



(11) **EP 1 840 210 A1**

(12) **EUROPEAN PATENT APPLICATION**
published in accordance with Art. 158(3) EPC

(43) Date of publication:
03.10.2007 Bulletin 2007/40

(21) Application number: **05819606.4**

(22) Date of filing: **16.12.2005**

(51) Int Cl.:
C12N 9/10 (2006.01) **C12N 15/54** (2006.01)
C12N 15/63 (2006.01) **C12N 15/82** (2006.01)
C12P 21/02 (2006.01)

(86) International application number:
PCT/CN2005/002219

(87) International publication number:
WO 2006/063533 (22.06.2006 Gazette 2006/25)

(84) Designated Contracting States:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI
SK TR**

(30) Priority: **16.12.2004 CN 200410102289**

(71) Applicant: **Peking University
Beijing 100871 (CN)**

(72) Inventors:
• **WANG, Yiping
Beijing 100871 (CN)**

• **SUN, Yicheng
Beijing 100871 (CN)**
• **LI, Yan
Beijing 100871 (CN)**

(74) Representative: **Jones, Elizabeth Louise et al
Frank B. Dehn & Co.
St Bride's House
10 Salisbury Square
London EC4Y 8JD (GB)**

(54) **RECONSTITUTION OF 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE ACTIVITY
BY FRAGMENT COMPLEMENTATION**

(57) The present invention relates to protein fragments of 5-enolpyruvylshikimate-3-phosphate synthase, which are selected from the protein fragment pairs of EPSPS, two such protein fragments can make up full length EPSPS and reconstitute EPSPS activities by complementation without help of any joint structure. The present invention also relates to nucleic acid molecules encoding the protein fragments, expression vectors and

cells comprising such nucleic acid molecules. The present invention also relates to methods for reconstituting EPSPS activities by using the fragments or the nucleic acid molecules or the expression vectors of the present invention, as well as methods for dividing the protein fragments of the present invention.

EP 1 840 210 A1

Description**TECHNICAL FIELD**

[0001] The present invention relates to a method for reconstitution of 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) activity by using fragment complementation. More particularly, the present invention relates to the reconstitution of the *Escherichia coli* EPSPS and the glyphosate-tolerant EPSPS mutants, as well as the reconstitution of *Pseudomonas putida* glyphosate-tolerant EPSPS.

TECHNICAL BACKGROUND

[0002] Along with the large scale planting of the transgenic plants, increasing attentions have been paid to the biological confinement of transgenic plants. It was observed that a gene encoding glyphosate-tolerant CP4 EPSPS could spread to crops 20 kilometers away from the transgenic plants, even into weeds, by Wu et al. Proc Natl Acad Sci U S A 2004, 101, (40), 14533-8. Thus it is very important to perform biological confinement towards transgenic plants for preventing superweeds and the like.

[0003] Modes of exogenous gene spreading of transgenic plants comprise principally the following: diffusion of the pollen from transgenic plants; transgenic plants acting as the recipients of pollen from wild sibling species to form hybrids; gene spreading which may be caused by the DNA of transgenic plants. Recently there are some methods which control the exogenous gene spreading of transgenic plants, e. g., (1) physical isolation, mainly distant isolation, so that the exogenous gene spreading by pollen is blocked; (2) genetic control, including: (a) male sterility; (b) genome incompatibility, that is, a specific exogenous gene is integrated into the crop genome which is incompatible with weeds; (c) maternal inheritance, which has achieved primary success in tobacco (Daniell et al., Nat Biotechnol 1998, 16, (4), 345-8), where exogenous genes are introduced into the chloroplasts of plants and subjected to maternal inheritance, thus no spreading into other species by diffusing pollen; (d) seed sterility; (e) transgenic mitigation (TM), where TM genes which are closely linked with target genes, advantageous or neutral to transgenic plants and disadvantageous to the growth of weeds (such as to prevent seeds from spilling out and to reduce the secondary dormancy of seeds) are used to prevent the production of superweeds.

[0004] A novel method for controlling the transgene flow of glyphosate-tolerant gene was reported by Ye, G. N. et al. Plant J 2001, 25, (3), 261-70 and Chin, H. G. et al. Proc Natl Acad Sci U S A 2003, 100, (8), 4510-5. In this method, a EPSPS gene was divided into two fragments which were then linked respectively with a gene expressing DanE intein and were co-expressed. A full length EPSPS was formed by self splicing of intein. The resulting *E. coli* or tobacco becomes tolerant to glyphosate. However, the intein-encoding gene itself was introduced into transgenic plants in this method, which might also lead to additional risks such as transgene flow.

[0005] Protein fragments derived from proteolytic cleavage or gene expression reconstitutes *in vivo* or *in vitro* into a complex with the same function as the intact protein. This is known as protein fragment complementation or protein reconstitution technology (e.g., Hakansson, M. et al. Curr Protein Pept Sci 2002, 3, (6), 629-42; Braun, M, et al. J Bacteriol 2003, 185, (18), 5508-18.). It is indicated by a research on fragment complementation of aminoacyl tRNA synthetase that most divisional sites of protein fragment for fragment complementation locate in non-conserved regions. That a protein can be reconstituted means the non-covalent interaction within this protein is quite specific, which makes the protein segments advantageous in forming a natural structure (Shiba, K. et al. Proc Natl Acad Sci USA 1992, 89, (5), 1880-4; Shiba, K. et al. J Biol Chem 1992, 267, (32), 22703-6). Furthermore, functional reconstitution of a protein means that the protein keeps a relative stable structure, even if the covalent bond is broken. Various non-covalent interactions (such as hydrogen bond, salt bridge and hydrophobic interactions) exist between two peptides for fragment complementation and the interactions are important for retaining stable structure of the protein (Nelson, K. E. et al., Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. Environ Microbiol 2002, 4, (12), 799-808).

[0006] We reconstituted active EPSPS *in vivo* and *in vitro* by such protein reconstitution. Glyphosate-tolerant EPSPS could also be reconstituted in this way. Fragment complementation of EPSPS can be used in transgenic plants to improve safety and reduce the possibility of superweed formation.

SUMMARY OF THE INVENTION

[0007] The present invention relates to 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) protein fragments, which are fragments selected from EPSPS fragment pairs. The pairs are comprised of two fragments that can connect to form full length EPSPS and reconstitute EPSPS activity by complementation without any joint structure. Preferably, EPSPS fragment pairs are derived from division of EPSPS within a joining region between folding units, between an α -helix and a β -sheet, or between two β -sheets, or within a β -sheet or a α -helix. More preferably, they are derived from

division of EPSPS within a joining region between folding units, e.g., in the joining regions between folding units 1 and 6, 2 and 6, 3 and 4, 4 and 5, and 3 and 5. More preferably, they are derived from division of EPSPS within one of the folding units 1, 2, 3, 4 or 5. For example, they are divided in the joining region between two β -sheets of folding unit 3, in the α -helix of folding unit 4, in the β -sheet of folding unit 2, between the α -helix and β -sheet of folding unit 1, in the joining region between the α -helix and β -sheet of folding unit 5, or in the joining region between two β -sheets of folding unit 5, or in the β -sheets of folding unit 5.

[0008] In the embodiments of the present invention, the EPSPS is wild type EPSPS or its active variants obtained by adding, deleting and/or replacing one or more amino acid residues. Preferably, the EPSPS is the wild type EPSPS of *E. coli* (its full length amino acid sequence and nucleotide sequence are known in the art and are shown in the Sequence Listing) or its glyphosate-tolerant EPSPS active variant. Preferably, the EPSPS is the glyphosate-tolerant EPSPS of *Pseudomonas putida*, such as of *P. putida* CGMCC 0739 (see Chinese Patent Application No. 02117991.3; its full length amino acid sequence and nucleotide sequence are shown in SEQ ID NO: 2 of the present application).

[0009] More specifically, the protein fragments of the present invention are preferably fragments from EPSPS fragment pairs of *E. coli* or of *P. putida* CGMCC 0739, wherein the fragment pairs of *E. coli* are selected from a group consisting of N67/C68, N85/C86, N104/C105, N154/C155, N182/C183, N184/C185, N218/C219, N224/C225, N227/C228, N259/C260, N298/C299, N371/C372, N376/C377, and N383/C384, (As used herein, the fragment pair N67/C68 is comprised of the N-terminal fragment N67 and the C-terminal fragment C68, wherein the N-terminal fragment N67 refers to the fragment from the N-terminus to the residue at position 67 in the EPSPS sequence, and the C-terminal fragment C68 refers to the fragment from the C-terminus to the residue at position 68 in the EPSPS sequence. This nomenclature applies to other fragment pairs.), and the fragment pairs of *P. putida* CGMCC 0739 are selected from a group consisting of N208/C209, N214/C215, N219/C220, N222/C223, and N224/C225. (As used herein, the fragment pair N208/C209 is comprised of the N-terminal fragment N208 and the C-terminal fragment C209, wherein the N-terminal fragment N208 refers to the fragment from the N-terminus to the residue at position 208 in the EPSPS sequence, and the C-terminal fragment C209 refers to the fragment from the C-terminus to the residue at position 209 in the EPSPS sequence. This nomenclature applies to other fragment pairs.)

[0010] The present invention further relates to a nucleic acid molecule encoding a protein fragment of the present invention, an expression vector carrying the nucleic acid molecule and a cell comprising the nucleic acid molecule or the expression vector. Preferably the cell is a plant cell. The present invention also relates to a transgenic plant comprising the nucleic acid molecule or the expression vector of the present invention or seeds thereof.

[0011] The present invention further relates to a method to reconstitute EPSPS, which includes reconstitution of EPSPS activity by using protein fragments or nucleic acid molecules or expression vectors of the present invention with no help of any joining structure.

[0012] The present invention further relates to a method of dividing EPSPS or its nucleic acid molecule with no help of any joining structure to obtain the protein fragment or the nucleic acid molecule of the present invention. The EPSPS is divided within a joining region between folding units, between an α -helix and a β -sheet, or between two β -sheets, or it is divided within a β -sheet or a α -helix. More preferably, it is divided within a joining region between folding units, e.g., in the joining region between folding units 1 and 6, 2 and 6, 3 and 4, 4 and 5, and 3 and 5. More preferably, it is divided within one of the folding units 1, 2, 3, 4 or 5. For example, it is divided in the joining region between two β -sheets of folding unit 3, in the α -helix of folding unit 4, in the β -sheet of folding unit 2, between the α -helix and β -sheet of folding unit 1, in the joining region between the α -helix and β -sheet of folding unit 5, or in the joining region between two β -sheets of folding unit 5, or in the β -sheets of folding unit 5. Even more preferably, the EPSPS is divided in a region selected from a group consisting of: the positions 67-68, 85-86, 104-105, 154-155, 182-183, 184-185, 218-219, N224-C225, N227-C228, 259-260, 298-299, 371-372, 376-377, or 383-384 in the EPSPS of *E. coli*, or the positions 208-209, 214-215, 219-220, 222-223, or 224-225 in the glyphosate-tolerant EPSPS of *P. putida* CGMCC 0739.

[0013] The present invention also relates to use of an EPSPS fragment or a method of reconstituting EPSPS activity or a method of dividing EPSPS according to the present invention in controlling the safety of transgenic plants.

DISCLOSURE OF THE INVENTION

[0014] Based on the structure of EPSPS, divisional sites are selected in the regions which might not influence the enzyme activity of EPSPS. Expression vectors are also constituted to express the divided EPSPS fragments. These fragments are then proved to be effective in complementing EPSPS activity in *E. coli*. Accordingly, the EPSPS fragments and the method for reconstitution of EPSPS as described herein could be used for the application in the field of biological confinement.

EPSPS and its structure

[0015] 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) is a key enzyme involved in aromatic amino acid bio-

synthesis - shikimic acid biosynthesis pathway in algae, higher plants, bacteria, fungi and apicomplexans of parasites. It catalyzes the yield of 5-enolpyruvylshikimate-3-phosphate (EPSP) from shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP).

[0016] As shown in Fig. 2, EPSPS is comprised of two domains, one includes the three symmetric folding units shown as Units 1, 2 and 6, and the other includes the three symmetric folding units shown as Units 3, 4 and 5. Each unit is comprised of two parallel α -helices and four β -sheets. See Stallings, W. C. et al. Proc Natl Acad Sci USA 1991, 88, (11), 5046-50.

[0017] EPSP synthase is known to form an "open" conformation when there is no substrate, but turns to a "close" conformation when it forms a complex crystal structure with S3P or with glyphosate+S3P (Schonbrunn, E. et al. Proc Natl Acad Sci U S A 2001, 98, (4), 1376-80). McDowell et al. (J Biomol NMR 2004, 28, (1), 11-29, 2004) modified the crystal structure of the complex of EPSP synthase, S3P and glyphosate by rotational-echo double-resonance NMR and obtained a three dimension structure which is consistent with the results from liquid NMR.

[0018] The substrate-binding sites and the catalytic sites of EPSP synthase are extensively investigated by methods such as chemical modification, gene mutation, and structural analysis (Schonbrunn, E. et al., *supra*; McDowell et al., *supra*; Anderson, K. S. et al. J Biol Chem 1990, 265, (10), 5567-72; Huynh, Q. K. et al. J Biol Chem 1988, 263, (24), 11636-9; Huynh, Q. K. et al. J Biol Chem 1988, 263, (2), 735-9; Padgett, S. R. et al. J Biol Chem 1988, 263, (4), 1798-802; Padgett, S. R. et al. Arch Biochem Biophys 1988, 266, (1), 254-62; Eschenburg, S. et al. J Biol Chem 2003, 278, (49), 49215-22; Mizyed, S. et al. Biochemistry 2003, 42, (23), 6986-95; Shuttlesworth, W. A. et al. Biochemistry 1994, 33, (23), 7062-8; Shuttlesworth, W. A. et al. Arch Biochem Biophys 1996, 334, (1), 37-42; Shuttlesworth, W. A. et al. Biochemistry 1999, 38, (1), 296-302; Stauffer, M. E. et al. Biochemistry 2001, 40, (13), 3951-7; Stauffer, M. E. et al. FEBS Lett 2001, 499, (1-2), 182-6; McDowell, L. M. et al. Biochemistry 2004, 43, (21), 6606-11).

[0019] Furthermore, it is indicated that Lys-22, Arg-124, Asp-313, Arg-344, Arg-386 and Lys-411 involve in PEP binding, Arg-27 involves in S3P binding, while Arg-100, Asp-242 and Asp-384 play important roles in conversion of the substrate-bound enzyme from an open conformation to a close one (Schonbrunn et al., Proteins 2000, 40, (2), 290-8; Biochemistry 2000, 39, (9), 2164-73; Proc Natl Acad Sci USA 2000, 97, (12), 6345-9).

Selection of divisional sites

[0020] The inventors selected divisional sites in the regions which might not influence the enzyme activity based on the structure of EPSPS.

[0021] When a divisional site locates between folding units, breakdown of covalent bond and insertion of methionine will generally not influence the formation of natural structure of the protein. The divided protein fragments form natural protein structure by non-covalent interaction within the protein and by various non-covalent interactions such as hydrogen bond, salt bridge and hydrophobic interactions between two peptides, in order to be able to reconstitute the protein function. When a divisional site locates between folding units, formation of secondary structure of the complementary fragments will generally not be influenced, thus the fragment complementation occurs more easily. In the embodiments, seven divisional sites are selected in the joining regions between folding units for *E. coli* EPSPS. Six pairs of the divided protein fragments are capable of complementing EPSPS activity with relatively high complementary activity. As for *P. putida* EPSPS, three divisional sites are selected in the joining regions between folding units and all the three pairs of fragments are able to complement EPSPS activity.

[0022] When a divisional site locates in a joining region between α -helices and/or β -sheets, breakdown of covalent bond will generally not influence the formation of the α -helix or β -sheet, and insertion of methionine will generally have no significant effect on the formation of natural structure, thus fragment complementation occurs easily. In the embodiments for *E. coli* EPSPS, six divisional sites are selected in the joining regions between α -helices and/or β -sheets, wherein 5 pairs of the fragments complement EPSPS activity; another divisional site is selected in the joining region between β -sheets, and the resulting fragments also complement EPSPS activity. As for *P. putida* EPSPS, one divisional site is selected in a β -sheet and the resulting fragments complement EPSPS activity.

[0023] When a divisional site is selected in an α -helix or a β -sheet, fragment complementation of divided fragments are achievable for reconstituting the activity as long as the α -helix or β -sheet is tolerant to the breakdown of covalent bond or to the insertion of methionine or is not important for the activity. In the embodiments for *E. coli* EPSPS, seven divisional sites are selected in an α -helix or a β -sheet, wherein 3 pairs of the fragments complement EPSPS activity. As for *P. putida* EPSPS, one divisional site is selected in a β -sheet and the relevant fragments complement EPSPS activity.

[0024] When EPSPS is divided between two domains, the N-terminal peptide N240 and the C-terminal peptide C241 of *E. coli* EPSPS are co-purified as a complex. However, the enzyme activity of this complex is very poor, which may be resulted from that there is only one pair of hydrophobic interaction regions in N240/C241 complex and the protein structure is thus unstable. As for *P. putida* EPSPS, three divisional sites are tested between two domains, while all the fragments are substantively not capable of complementing EPSPS activity. Accordingly, the EPSPS fragments obtained by dividing in the joining regions between domains could not reconstitute EPSPS activity.

[0025] In summary, the inventors successfully divide *E. coli* EPSPS at 14 sites, 6 of them locate in the joining regions between folding units, 3 of them locate in the joining regions between an α -helix and a β -sheet, 2 of them locate in the joining regions between β -sheets, 2 of them locate in β -sheets and 1 of them locates in an α -helix. By using these sites, the gene of *E. coli* EPSPS is divided into different N-terminal fragments and their relevant C-terminal fragments. Each N-terminal fragment and its relevant C-terminal fragment are carried separately on two compatible plasmids and are expressed. The resulting N-terminal protein fragment or its relevant C-terminal protein fragment alone does not exhibit EPSPS activity, while the association *in vivo* of the two fragments complement EPSPS activity.

[0026] Similarly, in another embodiment, the *P. putida* glyphosate-tolerant EPSPS is divided at 5 sites, wherein 3 sites locates in the joining regions between folding units, 1 in a joining region between an α -helix and a β -sheet, and 1 in a β -sheet. The resulting fragments also complement EPSPS activity.

[0027] The fact that EPSPS activity can be reconstituted by fragment complementation suggests the specific nature of the internal non-covalent interaction of this protein, which makes the fragments form a natural active structure exclusively and no other structure can be formed. In addition, the achievement of fragment complementation of the protein indicates that the protein keeps the structure stable even if covalent bonds are breakdown in some regions. That is to say, covalent bonds are not essential, non-covalent interactions such as hydrogen bond, salt bridge, and hydrophobic interaction are enough for the tight joint of two peptides. Hydrophobic interaction may be more important for EPSPS fragment complementation. The EPSPS fragment pairs have two potential hydrophobic-interaction regions which form two "hooks" to keep the protein stable.

[0028] A number of complexes are formed from fragment pairs in the present invention, but the amount of the purified proteins varies greatly. This indicates that the fragment complementation is always achievable *in vivo* but the form of the complexes from fragment pairs may be more or less difficult in different cases. The breakdown of a covalent bond has a great effect on protein structural stability in some regions, while in some other regions, this effect is less. Introduction of one methionine (the amino acid encoded by start codon) into a C-terminal peptide may have effects on stability of the protein structure, so that protein fragment complementation can not occur.

[0029] Complex N240/C241 of *E. coli* EPSPS is also co-purified in large amounts, but fails to complement the growth of *aroA* gene-mutant *E. coli* strain AB2829 on a restrictive medium, because the EPSPS activity of the Complex N240/C241 is much lower than that of the wild type full length EPSPS. The loss of activity possibly dues to two reasons: firstly, the divisional site of N240/C241 is very close to Asp242 which plays a very important role in the structural conversion upon substrate binding, and the introduction of one methionine into C241 may destroy the action of Asp242 which leads to the lose of enzyme activity; secondly, N240/C241 divides at a site between the two domains and the resulting complex has only one pair of hydrophobic interaction regions. Although the hydrophobic interaction between the regions is sufficient to make the two protein fragments form a complex, the structure of the complex may differ from that of the wild type protein in the absence of another pair of hydrophobic interaction region, thus the enzyme activity is lost. Accordingly, a divisional site should be selected based on the effect of the structural changes on enzyme activity.

Localization of EPSP synthase in plant cell

[0030] It was found that the aromatic amino acid biosynthesis takes place in chloroplast of plants (Bickel et al. Phytochemistry 17: 119-124, 1978). It was then demonstrated that the EPSP synthase locates in the endo-membrane of chloroplast in plants (Mousdale et al. Planta, 1987, 170:1~6; Plant Physiol., 1987, 83:229~231; J Biol Chem 1988 Oct 15;263 (29):15104-9; Mol Gen Genet 1994 Dec 1; 245(5):616-22; Mol Gen Genet 1993 Jun; 239(3):416-24). In *Petunia*, the molecular weight of precursor EPSP synthase (precursor EPSPS) is 55 kDa with the first 72 amino acid residues as a leading peptide, and the molecular weight of mature EPSPS is 48 kDa, which is formed by processing, cutting and transferring the precursor EPSPS after its entry into chloroplast (Della-Cioppa et al., Bio/Technology 1987(5): 579~584). A leading peptide plays an important role in entry of a precursor peptide into chloroplast. When a microbial EPSP synthase gene is introduced into plant cells, a nucleotide sequence encoding a plant leading peptide is added prior to the gene, otherwise the expressed microbial EPSPS, which has no natural leading peptide, can not enter into chloroplast.

Use

[0031] EPSPS fragment complementation technique of the present invention can be used in transgenic plants, in order to prevent the ecological risk caused by glyphosate-tolerant gene spreading. For instance, the N-terminal fragment of EPSPS is expressed in the nuclear chromosome of a plant, while the C-terminal fragment is expressed in the chloroplast. Expression of either of the EPSPS fragments alone does not bring about EPSPS activity, thus the gene spread into other plants lacks advantages for selection. The two EPSPS fragments complement in chloroplast to reconstitute EPSPS activity upon co-expression, which makes the plant become glyphosate tolerant. Alternatively, genes encoding two relevant EPSPS fragments may be inserted into different positions within the same chloroplast genome. In this case, the chance for both the two genes to be transferred into the chromosomes is significantly reduced and the chance for

gene spreading is greatly decreased. For the method to control the cell process in plants, please see WO2004/046359 and WO2004/046360.

DESCRIPTION OF THE DRAWINGS

[0032]

Figure 1 shows plasmid pKU2004.

Figure 2 shows the structure of EPSPS and selection of divisional sites. Fig. 2A shows the topological structure of EPSPS. Fig. 2B shows a structure unit of EPSPS.

Figure 3 shows the plasmid construction. A) A C-terminal fragment of EPSPS is subcloned next to a Tet promoter in pKU2100 vector. B) A N-terminal fragment of EPSPS is subcloned next to a Tet promoter in pACYC184 vector. C) A N-terminal part of the EPSPS encoding gene is subcloned next to a T7 promoter in a plasmid vector derived from pACYC184. D) A C-terminal part of EPSPS is constructed into pET28a vector, with a 6-His tag at the C-terminus.

Figure 4 shows the Western blot of EPSPS fragments. Cells expressing the indicated EPSPS are dissolved in SDS sample buffer. Equivalent amount of proteins (derived from ~0.8 µg wet cells) and the purified proteins are analyzed by SDS-PAGE and Western blot. Amounts of the following purified proteins are shown in the figure: N218/C219 (from ~150 ml wet cells), N227/C228 (from ~20 µg wet cells), N234/C235 (from ~500 µg wet cells), N240/C241 (from ~40 µg wet cells), N245/C246 (from ~500 µg wet cells) and pACYC184/pBR322 (from ~500 µg wet cells). The standard molecular weight is shown on the left side as in kilo-Dalton (kd).

Figure 5 shows the circular dichroism spectroscopy of EPSPS. Series 1 is for wild EPSPS and Series 2 is for N227/C228 complex.

Figure 6 shows the growth of *aroA* gene-deficient *E. coli* strain AB2829, which expresses different EPSPS, in the restrictive mediums (liquid M63 minimum medium in which glyphosate is added with the indicated concentrations).

Figure 7a shows SDS-PAGE of *E. coli* EPSPS. 1: N218; 2: C219; 3: EcEPSPS; 4: co-renatured N218 + c219; 5: N218+c219 (get together after individual renaturation); 6: standard molecular weight.

Fig. 7b shows Western blot of native-PAGE of *E. coli* EPSPS. 1: N218+C219 (get together after individual renaturation); 2: blank control; 3: co-renatured N218/C219; 4: blank control; 5: wild type EPSPS.

Examples

Example 1 Reconstitution of *E. coli* EPSPS

1 Materials and methods

1.1 Strains and plasmids

[0033] The strains and plasmids used in this study are listed in Table 1.

Table 1. Bacterial strains and plasmids used in this study

strain/plasmid	features	origin/reference
<u><i>E. coli</i> strain</u>		
DH5α	supE44ΔlacU169hsdR17recA1gyrA96thi-1re1A1	Hanahan D., J Mol Biol 1983;166:557-80
AB2829 B121(DE3)	<i>aroA</i> 354	Yale University stragene
BA ⁻	BL21 (DE3) □ <i>aroA</i> , Ap ^R	This study

(continued)

	Plasmid		
5	pUC18	ColE1, <i>lacZ'</i> , Ap ^R	Norrande <i>et al.</i> , 1983
	pBluscript-SK	ColE1, <i>lacZ'</i> , Ap ^R	Stratagene
	pET-28a	ColE, expression plasmid, Km ^R	Novagen
	pACYC184	Cm ^R	Chang, A. C. <i>et al.</i> 1978. J Bacteriol 134: 1141-56
10	pBR322	ColE1, Ap ^R	Schaeffer, F. <i>et al.</i> , 1982 EMBO J. 1, 99-105
	pKO3	Cm ^R	Link, A.J. <i>et al.</i> 1997. J. Bacteriology 179: 6228-6237
	pKU2005	pUC 18 derived plasmid with <i>aroA</i> _{<i>E.coli</i>} , Ap ^R	This study
	pKU2006	pACYC184 derived plasmid with <i>aroA</i> _{<i>E.coli</i>} , Cm ^R	This study
15	pKU2007	pACYC184 derived plasmid with <i>aroA</i> _{<i>E.coli</i>} -G96A, Cm ^R	This study
	pKU2008	pET-28a derived plasmid with <i>E.coli aroA</i> , Km ^R	This study
20	pKU2009	pET-28a derived plasmid with <i>E.coli aroA</i> , Km ^R	This study
	pKU2010	pET-28a derived plasmid with <i>E.coli aroA</i> -G96A, Km ^R	This study
	pKU2011	pET-28a derived plasmid with <i>E.coli aroA</i> -G96A, Km ^R	This study
25	pKU2100	pBR322 derived plasmid with <i>aroA</i> _{<i>E.coli</i>} , Ap ^R	This study
	pKU2101	pACYC184 derived plasmid with N218- <i>aroA</i> _{<i>E.coli</i>} , Cm ^R	This study
	pKU2102	pBR322 derived plasmid with C219- <i>aroA</i> _{<i>E.coli</i>} , Ap ^R	This study
30	pKU2107	pBR322 derived plasmid with <i>aroA</i> _{<i>E.coli</i>} -G96A, Ap ^R	This study
	pKU2110	pACYC184 derived plasmid with N245- <i>aroA</i> _{<i>E.coli</i>} , Cm ^R	This study
35	pKU2125	pBR322 derived plasmid with C228- <i>aroA</i> _{<i>E.coli</i>} , Ap ^R	This study
	pKU2126	pBR322 derived plasmid with C235- <i>aroA</i> _{<i>E.coli</i>} , Ap ^R	This study
	pKU2127	pET-28a derived plasmid with N218- <i>E.coli aroA</i> , Km ^R	This study
40	pKU2130	pBR322 derived plasmid with C246- <i>aroA</i> _{<i>E.coli</i>} , Ap ^R	This study
	pKU2137	pACYC184 derived plasmid with N234- <i>aroA</i> _{<i>E.coli</i>} , Cm ^R	This study
45	pKU2138	pACYC184 derived plasmid with N227- <i>aroA</i> _{<i>E.coli</i>} , Cm ^R	This study
	pKU2154	pBluscript-SK derived plasmid with <i>aroA</i> _{<i>E.coli</i>} , Ap ^R	This study
50	pKU2159	pET-28a derived plasmid with <i>E.coli aroA</i> , Km ^R	This study
	pKU2195	pBluscript-SK derived plasmid with C219- <i>aroA</i> _{<i>E.coli</i>} , Ap ^R	This study
	pKU2203	pET-28a derived plasmid with N218- <i>E.coli aroA</i> , Km ^R	This study
55	pKU2205	pET-28a derived plasmid with C218- <i>E.coli aroA</i> , Km ^R	This study

(continued)

	Plasmid		
5	pKU2223	pBluscript-SK derived plasmid with a part of the upstream sequence of <i>aroA_{E.coli}</i>	This study
	pKU2224	pBluscript-SK derived plasmid with <i>bla</i> gene	This study
	pKU2225	pBluscript-SK derived plasmid with the downstream sequence of <i>aroA_{E.coli}</i>	This study
10	pKU2227	pBluscript-SK derived plasmid with the upstream and downstream sequences of <i>aroA_{E.coli}</i>	This study
	pKU2228	pBluscript-SK derived plasmid with the upstream and downstream sequences of <i>aroA_{E.coli}</i> and the <i>bla</i> gene	This study
15	pKU2229	pK03 derived plasmid with the upstream and downstream sequences of <i>aroA_{E.coli}</i> and the <i>bla</i> gene	This study
	pKU2249	pET-28a derived plasmid with N245- <i>E.coli</i> <i>aroA</i> , Km ^R	This study
20	pKU2250	pET-28a derived plasmid with C246- <i>E.coli</i> <i>aroA</i> , Km ^R	This study
	pKU2262	pACYC184 derived plasmid with N240- <i>aroA_{E.coli}</i> , Cm ^R	This study
25	pKU2263	pBR322 derived plasmid with C241- <i>aroA_{E.coli}</i> , Ap ^R	This study
	pKU2266	pET-28a derived plasmid with N240- <i>E.coli</i> <i>aroA</i> , Km ^R	This study
	pKU2267	pACYC184 derived plasmid with T7 promoter and N218- <i>aroA_{E.coli}</i> , Cm ^R	This study
30	pKU2268	pACYC184 derived plasmid with T7 promoter and N240- <i>aroA_{E.coli}</i> , Cm ^R	This study
	pKU2269	pET-28a derived plasmid with N227- <i>E.coli</i> <i>aroA</i> , Km ^R	This study
35	pKU2274	pACYC184 derived plasmid with T7 promoter and N227- <i>aroA_{E.coli}</i> , Cm ^R	This study
	pKU2275	pACYC184 derived plasmid with T7 promoter and N234- <i>aroA_{E.coli}</i> , Cm ^R	This study
	pKU2276	pET-28a derived plasmid with C228- <i>E.coli</i> <i>aroA</i> , Km ^R	This study
40	pKU2277	pET-28a derived plasmid with C235- <i>E.coli</i> <i>aroA</i> , Km ^R	This study
	pKU2278	pET-28a derived plasmid with C241- <i>E.coli</i> <i>aroA</i> , Km ^R	This study
45	pKU2282	pET-28a derived plasmid with C246- <i>E.coli</i> <i>aroA</i> , Km ^R	This study
	pKU2283	pET-28a derived plasmid with C219- <i>E.coli</i> <i>aroA</i> , Km ^R	This study
	pKU2287	pACYC184 derived plasmid with T7 promoter and N240- <i>aroA_{E.coli}</i> , Cm ^R	This study
50	pKU2289	pACYC184 derived plasmid with T7 promoter and N245- <i>aroA_{E.coli}</i> , Cm ^R	This study

Ap, ampicillin; Cm, chloramphenicol; Km, kanamycin; ^R, resistance; Δ, deletion; ::, fusion.

55 1.2 Medium

[0034]

EP 1 840 210 A1

<u>LB media:</u>	amount per liter:
tryptone	10g
yeast extract	5g
NaCl	10g

Water was added up to 1 liter and pH was adjusted to around 7.0-7.5 with 2M NaOH. For solid media, 1.5% agar was added. The medium is autoclaved at 121°C, 15 pounds for 20 min before use.

[0035] Ap was added to a concentration of 50 µg/ml to select ampicillin resistant strains.

Km was added to a concentration of 25 µg/ml to select kanamycin resistant strains.

Cm was added to a concentration of 25 µg/ml to select chloromycetin resistant strains.

[0036] Restrictive M63 medium: 13.6g/L KH₂PO₄, 0.5mg/L FeSO₄·7H₂O, 20mM (NH₄)₂SO₄, 0.4% glucose, 1mM magnesium sulfate, 0.5mg/L vitamin B1.

1.3 Reagents

[0037] Restriction endonuclease, T₄DNA ligase, TaqDNA polymerase, DNA marker (Takara). Coomassie brilliant blue G250, enolpyruvate (sigma), shikimate-3-phosphate (a gift from Prof. Amrehin); HisTrap HP kit (Amersham Biosciences), goat anti-rabbit IgG (promega); all the other chemicals were reagents of analytical grade.

1.4 Genetic manipulation

[0038] Preparation of plasmid DNA, digestion with restriction endonuclease, ligation reaction, horizontal agarose electrophoresis in Tris-boric acid-EDTA buffer, polyacrylamide gel electrophoresis and Western blot and the like were carried out according to standard procedures (Molecular Cloning: A Laboratory Manual, 2nd edition. (Sambrook, Fritsch and Maniatis, eds.), Cold Spring Harbor Laboratory Press, 1989).

1.5 Plasmid construction

a. Construction of pKU2008, pKU2009

[0039] An *E. coli* *aroA* gene was amplified using primer 11: 5'-CGGGATCCAGGTCCGAAAAAACGCCGAC 3' and primer 12: 5'-CGGGATCCATGGAATCCCTGACGTTACA 3' and taking pKU2004 as a template, and was ligated into pET28a vector to obtain pKU2008. The EPSPS encoded by this plasmid was a fusion protein with His-tag at the N-terminus. pKU2008 was cleaved with *NcoI* and self-ligated to obtain pKU2009, which encodes an *E. coli* wild type EPSPS.

b. Construction of pKU2100 series plasmids (encoding C-terminal peptides of EPSPS)

[0040] PCR amplification was run using primer 13: 5'-TGAGTGACTGACTTTAAGAAGGAGATATAC3' and primer 14: 5'-CGGGATCCTCACTGATTTTCAATTTCAACAC 3' and taking pKU2009 as a template. The resulting product was digested with *Bam*HI and was ligated into the *Eco*RV and *Bam*HI sites of pBR322 to obtain plasmid pKU2100. Genes encoding C-terminal fragments of *E. coli* EPSPS were amplified taking pKU2009 as a template and using corresponding primers. The resulting products were ligated into the *NcoI* and *Bam*HI sites of pKU2100 to obtain plasmids pKU2102, pKU2125, pKU2126, pKU2130 and pKU2262, respectively. These plasmids encode C-terminal peptides of *E. coli* EPSPS (Table 1, Fig. 3).

c. Construction of pACYC184 series plasmids (encoding N-terminal peptides of EPSPS)

[0041] Genes encoding N-terminal fragments of *E. coli* EPSPS are amplified by PCR taking pKU2009 as a template and using corresponding primers. The resulting products were ligated into *Eco*RV and *Bam*HI sites of pACYC184 to obtain plasmids pKU2101, pKU2125, pKU2126, pKU2130 and pKU2263, respectively. These plasmids encode the amino acid sequences at the N-terminus of *E. coli* EPSPS (Table 1, Fig. 3).

d. Construction of pET28a series expression plasmids

[0042] The pBR322 series plasmid constructed in Step b and the pACYC184 series plasmid constructed in Step c were digested with *NcoI* and *Bam*HI respectively. Suitable fragments were recovered and ligated into pET28a vectors

to obtain pET28a series expression plasmids expressing N-terminal or C-terminal peptides of EPSPS (Table 1).

e. Construction of pACYC-T7 series expression plasmids

[0043] The pET28a plasmids expressing N-terminal peptides of EPSPS were digested with *Bgl*II and *Sal*I respectively. The desired fragments were recovered and ligated into *Bam*HI and *Sal*I sites of pACYC 184 vectors to obtain the pACYC-T7 series expression plasmids (Table 1, Fig. 3).

f. Construction of pET28a series expression plasmids (expressing fusion proteins of C-terminal peptides with His-tag)

[0044] Genes encoding different portions of *E. coli* EPSPS were amplified taking pKU2009 as a template and using corresponding primers. The resulting products were digested with *Xho*I and *Nco*I, and ligated into pET28a vectors to obtain pET28a series expression plasmids. The EPSPS encoded by the series plasmids was fused with 6 histidines at the C-terminus in order to be purified by a nickel column (Fig. 3).

g. Construction of plasmid pKU2229 for strain mutation

[0045] A ~ 600 bp fragment upstream of *aroA* gene in the chromosome of *E. coli* BL21 (DE3) was PCR amplified taking the chromosome as a template. The fragment was ligated into the *Bam*HI and *Hind*III sites of pBlueScript (stratagene) to obtain plasmid pKU2223. A *bla* gene of ~ 900 bp was PCR amplified taking plasmid pBlueScript as a template and was ligated into the *Hind*III and *Eco*RI sites of pBlueScript (stratagene) to obtain plasmid pKU2224.

[0046] A ~ 500 bp fragment downstream of *aroA* gene in the chromosome of *E. coli* BL21(DE3) was PCR amplified taking the chromosome as a template. The fragments were ligated into the *Eco*RI and *Sal*I sites of pBlueScript (stratagene) to obtain plasmid pKU2225. pKU2225 was digested by enzyme and the resulting fragments were ligated into the *Eco*RI and *Sal*I sites of pKU2223 to obtain plasmid pKU2227. pKU2224 was digested by enzyme and the resulting fragments were ligated into the *Hind*III and *Eco*RI sites of pKU2227 to obtain plasmid pKU2228. A fragment between the *Bam*HI site and *Sal*I site of pKU2228 was recovered by enzyme digestion and ligated into pKO3 to obtain plasmid pKU2229.

[0047] All the plasmids are confirmed of correct sequence by sequencing.

1.6 *Construction of mutant strain BA⁻*

[0048] The *aroA* gene of BL21 (DE3) was replaced with a *bla* gene, a gene conferring resistance to ampicillin, by using plasmid pKU2229 derived from pKO3. Specifically, pKU2229 was introduced into BL21 (DE3). A colony was picked up and diluted to spread on solid LB plate comprising antibiotics Ap and Cm, and was incubated overnight at 43°C. Since a pKO3 replicon was not able to originate replication normally at 43°C, a strain could grow on the plate containing antibiotics Ap and Cm at 43°C only if pKU2229 was recombined into the chromosome of the strain. One recombinant colony was picked up and diluted to spread onto the LB medium containing only Ap antibiotics and 5% sucrose, and was cultured overnight. Bacteria carrying a *Sac B* gene are not able to grow on the medium containing 5% sucrose because the protein encoded by the *Sac B* gene in pKO3 is toxic to bacterial cells due to degradation of sucrose. Accordingly, colonies grown on LB medium containing Ap antibiotics and 5% sucrose are the ones in which both *aroA* gene and *sacB* gene are deleted by a further homologous recombination. Several colonies were streaked on Cm plate and M63 plate to confirm the deletion of *aroA* gene. Total DNA was extracted from these colonies and PCR amplification was performed to confirm the replacement of *aroA* gene with *bla* gene.

1.7 *In vivo complementation*

[0049] The pKU2100 plasmid constructed as described above and corresponding pACYC1184 plasmid were introduced separately or together into the strain AB2829 deficient in *E. coli* *aroA* gene. Then, the strain was streaked on solid M63 restrictive medium and cultured overnight to determine the growth.

1.8 *Determination of Growth curve*

[0050] *aroA* gene-mutant *E. coli* strains AB 2829 carrying plasmid pKU2004, pKU2006 or pKU2007 were inoculated in liquid LB medium and incubated overnight. A sample of the culture was centrifuged at 4000 rpm for 3 min and re-suspended in 0.9% physiological saline. After a further centrifugation, the supernatant was discarded. The pellets were re-suspended in physiological saline and inoculated at an initial concentration of OD₆₀₀ 0.04 into a liquid media comprising 0, 50 or 100 mM glyphosate for overnight culture at 37°C. Absorption (OD₆₀₀) was measured at intervals.

1.9 Protein expression and purification

[0051] BA⁻ strain comprising the plasmid of interest was inoculated into liquid LB medium supplemented with corresponding antibiotics and incubated on a shaker at 37°C till OD₆₀₀ 0.75. IPTG was added to a final concentration of 0.5 mM into the medium and incubation was performed overnight at 15°C. A sample of the culture was centrifuged at 5000rpm, 4°C to collect cell pellets, which were then resuspended 10:1 by volume in Buffer A (50 mM Tris-HCl (pH 7.8), 0.4 mM DTT). The cells were broken by ultrasonication and centrifuged at 8000 rpm, 4°C for 60 min.

[0052] The supernatant was subjected to protein purification by using a HisTrap HP kit (Amersham Biosciences) according to the manufacturer's protocol. The purified proteins were concentrated in Buffer A using a Millipore Biomax membrane (10 kDa) and stored at 4°C.

1.10 Preparation of monoclonal antibody against EPSPS

[0053] Proteins encoded by pKU2008 in BL21 (DE3) strain were expressed and purified as described above, and then concentrated and resuspended in PBS buffer for use as an antigen to immunize rabbit. After four immunizations and one boost, serum was collected (this process was accomplished by Institute of Genetics, Chinese Academy of Science). The titre of the antibody was determined by ELISA prior to further use.

1.11 Polyacrylamide gel electrophoresis

[0054] SDS-PAGE (polyacrylamide gel electrophoresis) was run with the upper stacking gel of 5% and the lower separation gel of 16%. Native-PAGE was run with a gel of 10%. See, e.g. Sambrook *et al.*, *supra* for the preparation of gel and buffer and the method of electrophoresis.

1.12 Determination of Circular dichroism spectra (CD spectra)

[0055] The far ultraviolet- and near ultraviolet- CD spectra were determined on a Jobin Yvon CD6 connected with a thermostatic water-bath circulation controller. Each curve was obtained using average values from four measures. The measurement was performed at 8°C. A 1-cm pathlength cylinder quartz cuvette was used for the near ultraviolet CD spectra, and a 0.1-mm pathlength cylinder quartz cuvette was used for the far ultraviolet CD spectra.

1.13 Immunoblot of EPSPS fragment

[0056] Membrane transfer and immunoblot were carried out according to Sambrook *et al.*, *supra*. Briefly, the protein was transferred onto a cellulose nitrate membrane after subjected to 16% SDS-PAGE. The membrane was incubated with 1:2000 diluted rabbit monoclonal antibody. The resulting antigen-antibody complex was further reacted with a secondary antibody goat anti-rabbit IgG conjugated with horseradish peroxidase (promega) to form a complex. The color was developed upon the degradation of DAB (Sino-American Biotechnology Company) by horseradish peroxidase according to the manufacture's protocol.

1.14 Determination of EPSPS activity

Determination of Enzyme activity

[0057]

(1) 95 μl substrate (50 mmol/L HEPES buffer (pH 7.5), 1 mmol/L PEP, 1 mmol/L S3P) was pre-warmed at 28 °C in an incubator for 5 min.

(2) 5 μl enzyme was added and the reaction was kept in an incubator at 28 °C for 1-20 min (depends on the enzyme activity). After that, 800 μl MG/AM/NP was added and the reaction was mixed for 1 min.

(3) 100 μl 34% sodium citrate was added and the reaction was mixed for half an hour. The OD₆₆₀ was measured on a spectrometer.

1.15 In vitro complementation

[0058] The N-terminal and C-terminal peptides of EPSPS were expressed separately in BA⁻. The cells were collected by centrifugation and then broken by ultrasonication. Electrophoresis was run to see whether the proteins of interest precipitated mainly as inclusion bodies. The precipitate was washed first with Buffer B (50 mM Tris-HCl (pH 7.8), 1 mM

EDTA, 0.05% Triton-100) for three times, then with 1M NaCl for three times, then with Buffer C (50 mM Tris-HCl (pH 7.8), 1 mM EDTA, 1 M urea) for three times, and finally with distilled water for three times. The precipitate was then added into Buffer A containing 8M urea. After centrifugation, the supernatant was dialyzed against Buffer D (50 mM Tris-HCl (pH 7.8), 1 mM GSH, 0.5mM GSSG) with 2M urea overnight and then against Buffer D for 36 hours. Concentration was carried out using PEG12000. The renatured product was analyzed by electrophoresis, immune reaction and enzyme activity assay.

2. Result of fragment complementation of *E. coli* EPSPS

2.1 *In vivo* complementation

[0059] In order to test if EPSPS achieves fragment complementation after *in vivo* expression, genes encoding EPSPS N-terminal peptides were constructed into pACYC184, and genes encoding EPSPS C-terminal peptides were constructed into pBR322 (see the section *Materials and Methods* for the detailed construction method). Afterwards, these plasmids were introduced into *aroA* gene-mutant *E. coli* strain AB2829. EPSPS activity of the protein fragments encoded by these genes was determined by the growth of the strains on M63 restrictive medium. It was found that AB2829 strains carrying only N-terminal or C-terminal encoding genes are not able to grow on the M63 medium.

[0060] The pACYC184 and pBR322 plasmids are able to be expressed in the same bacterial cell since they have replicons p15A and ColEI respectively. In this regard, a pACYC184 plasmid was co-introduced into *aroA* gene-mutant *E. coli* strain AB2829 together with a corresponding pKU2100 plasmid, wherein the N-terminal peptide encoded by the pACYC184 plasmid and C-terminal peptide encoded by the pKU2100 plasmid are the two fragments divided at the same site from EPSPS. The resulting strains are streaked on M63 restrictive medium, and cultured at 37°C. 16 hours later, AB2829 strains carrying N218/C219-encoding plasmids (pKU2101/pKU2102) or N227/C228-encoding plasmids (pKU2125/pKU2138) grow on the restrictive medium, while AB2829 strains carrying other three pairs of plasmids pKU2126/ pKU2137(N234/C235), pKU2162/pKU2163(N240/C241) or pKU2110/pKU2130 (N245/C246) fail to grow on the M63 medium. This indicated that EPSPS fragment pairs N218/C219 and N227/C228 complement EPSPS activity upon their *in vivo* expression, while the other three pairs of fragments not.

2.2 Purification and determination of complex of EPSPS fragments

[0061] In order to determine whether the recovery of EPSPS activity is due to *in vivo* recombination of genes encoding EPSPS fragments which leads to a full length EPSPS, the protein fragments were purified and tested *in vitro* for enzyme activity and sizes to confirm that the recovery of EPSPS activity is due to fragment complementation. For this, plasmids are constructed for expression of EPSPS fragments under the control of T7 promoter, wherein the N-terminal fragments are constructed on pACYC184 vectors and the C-terminal fragments fused with 6 histidines are constructed on pET28a vectors (see the context above for the method of plasmid construction). If pairs of two fragments are able to form a natural structure, the N-terminal peptide will associate with the corresponding C-terminal peptide to form a complex which could be purified by attachment onto a nickel column.

[0062] In order to eliminate the effect of EPSPS encoded by *aroA* gene in the bacterial chromosome, a strain BA was obtained in which the *aroA* gene in *E. coli* BL21 was replaced by a *bla* gene encoding β -lactamase. The pairs of plasmids were introduced into a BA strain for co-expression. The cells were streaked on a M63 plate containing IPTG and cultured overnight for *in vivo* complementation. The result was consistent with the above description: BA⁻ cells carrying genes for N218/C219 or N227/C228 fragments grow on the M63 restrictive medium; BA⁻ cells carrying genes of the other three fragment pairs fails to grow on the M63 restrictive medium.

[0063] The N-terminal peptides and the C-terminal peptides of EPSPS were expressed separately or together in BA⁻ cells and their expression are confirmed by SDS-PAGE. The co-expressed proteins were purified using a HisTrap HP kit, and probed by using an EPSPS antibody on SDS-PAGE (see the section *Materials and methods* for the detailed procedure). As shown in Fig. 4, the purified proteins comprise both the C-terminal and the corresponding N-terminal peptides and the sizes of these protein fragments are the same as expected. This indicated that the EPSPS fragments co-expressed *in vivo* are purified as complexes due to complementation. It was also obvious in the figures that the co-expressed fragments are not linked together to be a full length protein before or after purification, which indicates that no recombination and thus no full length EPSPS gene occurred *in vivo*. It was also shown by the activity assay that all of the purified proteins had EPSPS activity. In summary, EPSPS activity is recovered by fragment complementation of peptides expressed respectively *in vivo*.

[0064] The expression amount of N- and C-terminal peptides of EPSPS may vary. For example, the expression amount of N227 was much more than that of C228, but the ratio of N- and C-terminal fragments after purification was 1:1 in almost all cases. As shown in the figures, the amounts of N234/C235 and N245/C246 after purification were much less than that of the other three co-purified EPSPS. May be this is why they fail to complement EPSPS activity *in vivo* to

make AB2829 grow on the restrictive medium. Although the amount of N240/C241 complex after purification was even more than that of N218/C219, it fails to support the growth of AB2829 on the restrictive medium because of the poor EPSPS activity.

2.3 Determination of enzyme activity

[0065] N218/C219 and N227/C228, which are purified with HisTrap HP kit, and a full length EPSPS were further purified on a Sephadex-G75 column. As a result, N218/C219, N227/C228 and the full length EPSPS shared the same elution peak, indicating that they were substantially the same in terms of molecular weight and structure. The enzyme activities and the K_m values to substrates of these three proteins were measured. As shown in Table 2, the enzyme activities of EPSPS complexes N218/C219 and N227/C228 after fragment complementation were 70% and 64% of that of full length EPSPS, respectively, while the affinity to substrates changed little, which indicated that the structure of the reconstituted EPSPS was relatively stable.

Table 2. Enzymatic properties of EPSPS and the reconstituted EPSPS

Enzyme (μM)	specific activity ^b ($\text{mol min}^{-1} \text{mg}^{-1}$)	K_m [PEP] ^c (μM)	K_m [S3P] ^d
EPSPS	44 ± 2	14 ± 3	45 ± 1
218/219	31 ± 3	18 ± 3	49 ± 3
227/228	28 ± 4	19 ± 3	49 ± 6

^a Results from two individual experiments, each of which includes triplicates.

^b For determination of the enzyme specific activity, the concentrations of PEP and S3P were both 1.0 mM.

^c For determination of K_m [PEP], the concentration of S3P was kept as 1 mM while the concentration of PEP varied from 50 to 200 μM .

^d For determination of K_m [S3P], the concentration of PEP was kept as 1 mM while the concentration of S3P varied from 50 to 200 μM .

2.4 Circular dichroism spectra (CD spectra) analysis

[0066] In order to confirm whether the structure of the reconstituted EPSPS was the same as wild type EPSPS, circular dichroism spectroscopy was carried out for EPSPS and the complex N227/C228. There was no significant difference in far ultraviolet CD spectra between the EPSPS and N227/C228 (Fig. 5), which indicated that their secondary structures were almost the same. However, there was difference in near ultraviolet CD spectra (Fig. 5), which indicated that there was difference in structure between the reconstituted EPSPS and the wild type EPSPS.

2.5 *In vivo* complementation of mutant EPSPS of *E. coli*

[0067] In order to test whether the mutated EPSPS fragments perform functional complementation *in vivo*, plasmid pKU2105 and pKU2333 were constructed, which encode the N-terminal fragments N218-G96A and N227-G96A of *E. coli* EPSPS, respectively. They were introduced into *aroA*-mutant *E. coli* strain AB2829 (*aroA*⁻, from Yale University) together with their corresponding plasmids encoding C219 and C228 respectively. The growth was detected in the M63 restrictive media with different concentrations of glyphosate. As shown in Fig. 6, all the AB2829 strains grew well in M63 medium with no glyphosate without significant difference. In M63 medium containing 50 mM glyphosate, strains expressing the mutant full-length EPSPS (EPSPS-G96A) or the mutant complementary fragments (N227-G96A/C228 and N218-G96A/C219) grew well, while strains expressing wild type EPSPS or wild type complementary fragments were inhibited from growth. In the M63 medium containing 100 mM glyphosate, strains expressing EPSPS-G96A or N227-G96A/C228 still grew well while strains expressing N218-G96A/C219 did not grow as well as they did in the M63 medium containing 50 mM glyphosate, which might be because that the N227-G96A/C228 complex is more easy to be formed *in vivo*.

2.6 *In vitro* reconstitution of EPSPS

[0068] In case that protein fragments form active complex by complementation *in vivo*, they are expected to form active complex *in vitro*. To verify this, EPSPS N-terminal peptide N218 and C-terminal peptide C219 were expressed separately. Inclusion bodies were formed in the presence of rich EPSPS fragments being expressed, and were subjected to purification and renaturation (see the section *Materials and methods* for the procedure). Result of electrophoresis

after renaturation was shown in Fig. 7. As shown in Fig. 7a, each EPSPS fragment presented as expected and no full length EPSPS was found. *In vitro* co-renaturation of N218 and C219 leads to a band in native-PAGE which were consistent with that of wild type EPSPS, while the mixture of separately renatured N218 and C219 did not show this band. This indicates that a structure similar to that of wild type EPSPS could be formed upon *in vitro* co-renaturation of EPSPS. The activity assay also showed that complementation of EPSPS activity is achievable upon co-renaturation of N218/C219, but not upon mixture of separately renatured N218 and C219.

3. Correlations between fragment complementation and structure of EPSPS

3.1 The strains and plasmids used are show in the table below.

[0069]

Table 3. Bacteria strains and plasmids used in this study

strain/plasmid	features	Origin/Reference
<u><i>E. coli</i> strain</u>		
DH5 α	supE44 Δ lacU169hsdR17recA1gyrA96thi-1re1A1	Hanahan D., J Mol Biol 1983;166:557-80
AB2829	aroA354	Yale University
<u>plasmid</u>		
pUC18	ColE1, <i>lacZ'</i> , Ap ^R	Norrandner <i>et al.</i> , 1983
pBluscript-SK	ColE1, <i>lacZ'</i> , Ap ^R	Stratagene
pET-28a	ColE1, expression plasmid, Km ^R	Novagen
pACYC184	Cm ^R	Chang, A. C. et al. 1978. J Bacteriol 134:1141-56
pBR322	ColE1, Ap ^R	Schaeffer, F. et al., 1982 EMBO J. 1, 99-105
pKU2005	pUC 18 derived plasmid with <i>aroA</i> _{<i>E. coli</i>} , Ap ^R	This study
pKU2006	pACYC184 derived plasmid with <i>aroA</i> _{<i>E. coli</i>} , Cm ^R	This study
pKU2007	pACYC184 derived plasmid with <i>aroA</i> _{<i>E. coli</i>} -G96A, Cm ^R	This study This study
pKU2008	pET-28a derived plasmid with <i>E. coli aroA</i> , Km ^R	This study
pKU2009	pET-28a derived plasmid with <i>E. coli aroA</i> , Km ^R	This study
pKU2010	pET-28a derived plasmid with <i>E. coli aroA</i> -G96A, Km ^R	This study
pKU2011	pET-28a derived plasmid with <i>E. coli aroA</i> -G96A, Km ^R	This study This study
pKU2100	pBR322 derived plasmid with <i>aroA</i> _{<i>E. coli</i>} , Ap ^R	This study
pKU2101	pACYC184 derived plasmid with N218- <i>aroA</i> _{<i>E. coli</i>} , Cm ^R	This study This study
pKU2102	pBR322 derived plasmid with C219- <i>aroA</i> _{<i>E. coli</i>} , Ap ^R	This study
pKU2103	pACYC184 derived plasmid with N238- <i>aroA</i> _{<i>E. coli</i>} , Cm ^R	This study
pKU2104	pBR322 derived plasmid with C239- <i>aroA</i> _{<i>E. coli</i>} , Ap ^R	This study
pKU2110	pACYC184 derived plasmid with N245- <i>aroA</i> _{<i>E. coli</i>} , Cm ^R	This study

(continued)

	plasmid	
5	pKU2111	pACYC184 derived plasmid with N259- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
	pKU2124	pBR322 derived plasmid with C225- <i>aroA</i> _{<i>E. coli</i>} Ap ^R This study
	pKU2125	pBR322 derived plasmid with C228- <i>aroA</i> _{<i>E. coli</i>} Ap ^R This study
10	pKU2126	pBR322 derived plasmid with C235- <i>aroA</i> _{<i>E. coli</i>} Ap ^R This study
	pKU2129	pACYC184 derived plasmid with N31- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
15	pKU2130	pBR322 derived plasmid with C246- <i>aroA</i> _{<i>E. coli</i>} Ap ^R This study
	pKU2131	pBR322 derived plasmid with C260- <i>aroA</i> _{<i>E. coli</i>} Ap ^R This study
	pKU2135	pBR322 derived plasmid with C299- <i>aroA</i> _{<i>E. coli</i>} Ap ^R This study
20	pKU2136	pACYC184 derived plasmid with N298- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
	pKU2137	pACYC184 derived plasmid with N234- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
25	pKU2138	pACYC184 derived plasmid with N227- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
	pKU2139	pACYC184 derived plasmid with N224- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
30	pKU2148	pACYC184 derived plasmid with N165- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
	pKU2149	pBR322 derived plasmid with C166- <i>aroA</i> _{<i>E. coli</i>} Ap ^R This study
	pKU2150	pACYC184 derived plasmid with N371- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
35	pKU2151	pBR322 derived plasmid with C372- <i>aroA</i> _{<i>E. coli</i>} Ap ^R This study
	pKU2290	pACYC184 derived plasmid with N67- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
40	pKU2291	pACYC184 derived plasmid with N73- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
	pKU2292	pACYC184 derived plasmid with N84- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
45	pKU2293	pACYC184 derived plasmid with N104- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
	pKU2294	pACYC184 derived plasmid with N154- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
	pKU2295	pACYC184 derived plasmid with N182- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
50	pKU2296	pACYC184 derived plasmid with N184- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
	pKU2297	pACYC184 derived plasmid with N376- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
55	pKU2298	pACYC184 derived plasmid with N383- <i>aroA</i> _{<i>E. coli</i>} Cm ^R This study
	pKU2299	pBR322 derived plasmid with C68- <i>aroA</i> _{<i>E. coli</i>} Ap ^R This study

(continued)

<u>plasmid</u>		
5	pKU2300	pBR322 derived plasmid with C74- <i>aroA</i> _{<i>E. coli</i>} ^{Ap^R} This study
	pKU2301	pBR322 derived plasmid with C86- <i>aroA</i> _{<i>E. coli</i>} ^{Ap^R} This study
	pKU2302	pBR322 derived plasmid with C 105- <i>aroA</i> _{<i>E. coli</i>} ^{Ap^R} This study
10	pKU2303	pBR322 derived plasmid with C155- <i>aroA</i> _{<i>E. coli</i>} ^{Ap^R} This study
	pKU2304	pBR322 derived plasmid with C183- <i>aroA</i> _{<i>E. coli</i>} ^{Ap^R} This study
15	pKU2305	pBR322 derived plasmid with C185- <i>aroA</i> _{<i>E. coli</i>} ^{Ap^R} This study
	pKU2306	pBR322 derived plasmid with C377- <i>aroA</i> _{<i>E. coli</i>} ^{Ap^R} This study
20	pKU2307	pBR322 derived plasmid with C384- <i>aroA</i> _{<i>E. coli</i>} ^{Ap^R} This study
Ap, ampicillin; Cm, chloromycetin; ^R , resistance.		

3.2 Medium

[0070] See Section 1.2 of this example.

3.3 Reagent

[0071] Restriction endonuclease, T₄DNA ligase, DNA polymerase, and DNA marker were from Takara. All the other chemicals were reagents of analytical grade.

3.4 Genetic manipulation

[0072] Preparation of plasmid DNA, digestion with restriction endonuclease, ligation reaction and horizontal agarose electrophoresis in Tris-boric acid-EDTA buffer were carried out according to standard procedures (Maniatis *et al.*, 1982).

3.5 Plasmid construction

a. Construction of pBR322 series plasmid (encoding C-terminal peptides of EPSPS)

[0073] The genes encoding C-terminal peptides of *E. coli* EPSPS were amplified taking pKU2009 (*E. coli aroA* gene) as a template. The resulting products were ligated into the *Nco*I and *Bam*HI sites of pKU2100 to obtain the series plasmids encoding C-terminal peptides of EPSPS (Table 3).

b. Construction of pACYC184 series plasmid (encoding N-terminal peptides of EPSPS)

[0074] The genes encoding N-terminal peptides of *E. coli* EPSPS were amplified taking pKU2009 or pGMO as a template. The resulting products were ligated into the *Eco*RV and *Bam*HI sites of pACYC184 to obtain the series plasmids encoding N-terminal peptides of EPSPS (Table 3).

[0075] All the plasmids constructed above are confirmed of correct sequence by sequencing.

3.6 In vivo complementation

[0076] The pBR322 plasmids constructed in section 3.5 a and their relevant pACYC184 plasmids constructed in section 3.5 b were introduced respectively or together into *E. coli* strain AB2829 deficient in *aroA* gene. Then the strains were streaked on solid M63 restrictive medium and cultured overnight to test the growth.

3.7 Results

[0077] 21 divisional sites were designed in total based on the structure of *E. coli* EPSPS and were investigated for EPSPS fragment complementation. 3 divisional sites locate in α -helix, wherein neither of N31/C32 or N245/C246 complements the growth of *aroA* gene mutant *E. coli* strain AB2829 on the restrictive medium, while the complementation of N105/C106 was poor; 3 divisional sites locate in β -sheet, wherein neither of N73/C74 or N238/C239 complement the EPSP enzyme activity, while N224/C225 complement the growth of AB2829 on the restrictive medium; 5 divisional sites locate in the joining regions between β -sheets or between α -helices, wherein only N165/C166 fails to complement the growth of AB2829 on the M63 restrictive medium, while the rest four pairs of EPSPS fragments complement EPSPS activity; 6 divisional sites locate between two folding units, wherein only N234/C235 fails to complement the EPSPS activity, while the rest five pairs of fragments complement well on the growth of AB2829 on the M63 restrictive medium.

Table 4 Fragment complementation of *E. coli* EPSPS

EPSPS fragment		region where the divisional complementation condition site locates			
	N-terminal	C-terminal	N-terminal	C-terminal	N-terminal + C-terminal
	N31	C32	α -helix	-	-
	N67	C68	between two β -sheets	-	+
	N73	C74	β -sheet	-	-
	N85	C86	between folding unit 3 and 4	-	+++
	N104	C105	α -helix	-	+/-
	N154	C155	between folding unit 4 and 5	-	+++
	N165	C166	between α -helix and β -sheet	-	-
	N182	C183	between α -helix and β -sheet	-	+
	N184	C185	between α -helix and β -sheet	-	++
	N218	C219	between two β -sheets	-	++
	N224	C225	β -sheet	-	+
	N227	C228	between folding unit 3 and 5	-	+++
	N234	C235	between folding unit 3 and 5	-	-
	N238	C239	β -sheet	-	-
	N240	C241	between two domains	-	-
	N245	C246	α -helix	-	-
	N259	C260	β -sheet	-	+
	N298	C299	between folding unit 2 and 6	-	+++
	N371	C372	between folding unit 1 and 6	-	+++
	N376	C377	between folding unit 1 and 6	-	+++
	N383	C384	between α -helix and β -sheet	-	+

-, fail to grow; +, be able to grow.

Example 2 Fragment complementation of *P. putida* glyphosate-tolerant EPSPS

[0078] Fragment complementation experiment was carried out by using the glyphosate-tolerant EPSPS gene of *P. putida* CGMCC 0739. The plasmids, media, strains and the protocols are the same as in Example 1. The reconstituted enzyme activity was tested by growth on the medium containing 100 mM glyphosate.

Table 5 Fragment complementation of *P. putida* glyphosate-tolerant EPSPS

EPSPS fragment	region where the divisional site locates	complementation condition
N208/C209	between α -helix and β -sheet	+
N214/C215	in β -sheet	+
N219/C220	between folding units	++
N222/C223	between folding units	+
N224/C225	between folding units	++
N233/C234	between two domains	-
N234/C234	between two domains	-
N236/C237	between two domains	-

[0079] Primers used are listed as follows:

N-terminal forward primer:

ppN5': 5' -TGA GTG ACT GAA AGT GAA AGT AAC AAT ACA G-3' N-terminal reversed primer is described below.

C-terminal reverse primer:

ppC3'(BamHI): 5'- CGG GAT CCC TTC TTC GGA CAA TGA CAG AC-3' C-terminal reverse primer is described below.

N208/C209 +

ppN2083' (BamHI): 5'- CGG GAT CCT CAG GGA GTC TTC AAA CCA AAC C-3'
ppC2095'(NcoI): 5'- CAT GCC ATG GAG AAT CGAAAC TAT GAA G-3'

N214/C215 +

ppN2143' (BamHI): 5'-CGG GAT CCT CAT TCA TAG TTT CGA TTC TCG G-3'
ppC2155'(NcoI) : 5'-CAT GCC ATG GAG TTT TAT TTC AAA GCC GG-3'

N219/C220+

ppN2193' (BamHI): 5'-CGG GAT CCT CAT TTG AAA TAA AAC TCT TCA TAG-3'
ppC2205' (NcoI): 5'- CAT GCC ATG GCC GGG AAT GTA TAT GAT GAA AC-3'

N222/C223 +

ppN2223' (BamHI): 5'- CGG GAT CCT CAA TTC CCG GCT TTG AAA TAA AAC-3'
ppC2235'(NcoI) : 5'- CAT GCC ATG GTA TAT GAT GAA ACG AAA ATG-3'

N224/C225 + +

ppN2243'(BamHI): 5'- CGG GAT CCT CAA TAT ACA TTC CCG GCT TTG-3'
ppC2255' (NcoI): 5'- CAT GCC ATG GAT GAA ACG AAA ATG CAA CG-3'

N233/C234

ppN2333' (BamHI): 5'- CGG GAT CCT CAG GTG TAT CGT TGC ATT TTC G-3'
ppC2345'(NcoI): 5'- CAT GCC ATG GTA GAA GGC GAC TGG AGC G-3'

N234/C235

EP 1 840 210 A1

ppN2343' (BamHI): 5'- CGG GAT CCT CAT ACG GTG TAT CGT TGC ATT TTC-3'
ppC2355' (NcoI): 5'- CAT GCC ATG GAA GGC GAC TGG AGC GGT GG-3'

N236/C237

5

ppN2363' (BamHI): 5'- CGG GAT CCT CAG CCT TCT ACG GTG TAT CGT TG-3'
ppC2375'(NcoI) : 5'- CAT GCC ATG GAC TGG AGC GGT GGT GCT TT-3'

10

15

20

25

30

35

40

45

50

55

Sequence Listing

5 <110> Peking University
Wang, Yiping
Sun, Yicheng
Li, Yan

10 <120> Reconstitution of 5-enolpyruvylshikimate-3-phosphate synthase activity by
fragment complementation

<130>

15 <141>

<160> 79

20 <170> PatentIn version 3.1

<210> 1
<211> 1501
<212> DNA
25 <213> Pseudomonas putida P.P4G-1

<400> 1
agagtttgat catggctcag attgaacgct ggccggcaggc ctaacacatg caagtcgagc 60
30 ggatgagaag agcttgctct tcgattcagc ggccggacggg tgagtaatgc ctaggaatct 120
gcctggtagt gggggacaac gtttcgaaag gaacgctaata accgcatacg tcctacggga 180
gaaagcaggg gaccttcggg ccttcgcgta tcagatgagc ctaggtcgga ttagctagtt 240
35 ggtgaggtaa tggctcacca aggcgacgat ccgtaactgg tctgagagga tgatcagtc 300
cactggaact gagacacggt ccagactcct acgggaggca gcagtgggga atattggaca 360
atgggcgaaa gcctgatcca gccatgccgc gtgtgtgaag aaggtcttcg gattgtaaag 420
40 cactttaagt tgggaggaag ggcattaacc taatacgta gtgttttgac gttaccgaca 480
gaataagcac cggctaactc tgtgccagca gccgcggtaa tacagagggt gcaagcgta 540
45 atcggaaatta ctgggcgtaa agcgcgcgta ggtggtttgt taagttggat gtgaaagccc 600
cgggctcaac ctgggaactg tatccaaaac tggcaagcta gactacgga gaggggtggtg 660
gaatttcctg tgtagcgggtg aaatgcgtag atataggaag gaacaccagt ggcgaaggcg 720
50 accacctgga ctgatactga cactgagggtg cgaaagcgtg gggagcaaac aggattagat 780
accctggtag tccacgccgt aaacgatgtc aactagccgt tggaatcctt gagattttag 840
tggcgcagct aacgcattaa gttgaccgcc tggggagtag gccgcgaagg ttaaaactca 900
55 aatgaattga cgggggcccc cacaagcggg ggagcatgtg gtttaattcg aagcaacgcg 960

EP 1 840 210 A1

	aagaacctta ccaggccttg acatgcagag aacittccag agatggattg gtgccttcgg	1020
5	gaactctgac acagggtctg catggctgtc gtcagctcgt gtcgtgagat gttgggttaa	1080
	gtcccgtaac gagcgcaacc cttgtcctta gttaccagca cgtaatgggtg ggcaactctaa	1140
	ggagactgcc ggtgacaaac cggaggaagg tggggatggc gtcaagtcac catggccctt	1200
10	acggcctggg ctacacacgt gctacaatgg tcggtacaga gggttgccaa gccgcgaggt	1260
	ggagctaata tcacaaaacc gatcgtagtc cggatcgag tctgcaactc gactgcgtga	1320
15	agtcggaata gctagtaata gcgaatcaga atgtcgcggt gaatacgttc ccgggccttg	1380
	tacacaccgc ccgtcacacc atgggagtggt gttgcaccag aagtagctag tctaaccctc	1440
	gggaggacgg ttaccacggt gtgattcatg actgggtga agtcgtaaca aggtaaccgt	1500
20	a	1501
	<210> 2	
	<211> 1914	
	<212> DNA	
25	<213> Pseudomonas putida P.P4G-1	
	<220>	
	<221> CDS	
	<222> (508).. (1800)	
	<223>	
30	<400> 2	
	gatcataaaa catgcttgta taaaggatgc tgccatgttc cgtgaactgg aagcgaacaa	60
35	tcttgcggta tatcagaaaa agccaaagct gattgcagtg cttcttcagc gtaatgctca	120
	gttaaaagcg aaggttggtc aggaggatga gttcgaaaag tcggtaaggc gtttgttgaa	180
	ctttggtcat acattggggc atgccatcga aaatgaatat gcgttgatgc atggccatgc	240
40	ggttgctata ggaatgacat acgcgtgtca tatttctgag caattgtctg gattcaaaca	300
	aacaaatcgc gtggtagaag tgttggaaca atatgggtta ccgacttata tggcattcga	360
	tagggaaaag gcttttaata tgttgaaaat ggacaagaag cgtgaaaaaa aggaaatgaa	420
45	ctatgtgttg ctggaaaaag tagggaaggg agtgggtgaag agtattccac tggttcaatt	480
	agaaaaaatc attcaagcat taccaaa gtg aaa gta aca ata cag ccc gga gat	534
	Met Lys Val Thr Ile Gln Pro Gly Asp	
50	1 5	
	ctg act gga att atc cag tca ccc gct tca aaa agt tcg atg cag cga	582
	Leu Thr Gly Ile Ile Gln Ser Pro Ala Ser Lys Ser Ser Met Gln Arg	
	10 15 20 25	
55	gct tgt gct gct gca ctg gtt gca aaa gga ata agt gag atc att aat	630
	Ala Cys Ala Ala Ala Leu Val Ala Lys Gly Ile Ser Glu Ile Ile Asn	

EP 1 840 210 A1

	30	35	40	
5	ccc ggt cat agc aat gat gat aaa gct gcc agg gat att gta agc cgg Pro Gly His Ser Asn Asp Asp Lys Ala Ala Arg Asp Ile Val Ser Arg 45 50 55	678		
10	ctt ggt gcc agg ctt gaa gat cag cct gat ggt tct ttg cag ata aca Leu Gly Ala Arg Leu Glu Asp Gln Pro Asp Gly Ser Leu Gln Ile Thr 60 65 70	726		
15	agt gaa ggc gta aaa cct gtc gct cct ttt att gac tgc ggt gaa tct Ser Glu Gly Val Lys Pro Val Ala Pro Phe Ile Asp Cys Gly Glu Ser 75 80 85	774		
20	ggt tta agt atc cgg atg ttt act ccg att gtt gcg ttg agt aaa gaa Gly Leu Ser Ile Arg Met Phe Thr Pro Ile Val Ala Leu Ser Lys Glu 90 95 100 105	822		
25	gag gtg acg atc aaa gga tct gga agc ctt gtt aca aga cca atg gat Glu Val Thr Ile Lys Gly Ser Gly Ser Leu Val Thr Arg Pro Met Asp 110 115 120	870		
30	ttc ttt gat gaa att ctt ccg cat ctc ggt gta aaa gtt aaa tct aac Phe Phe Asp Glu Ile Leu Pro His Leu Gly Val Lys Val Lys Ser Asn 125 130 135	918		
35	cag ggt aaa ttg cct ctc gtt ata cag ggg cca ttg aaa cca gca gac Gln Gly Lys Leu Pro Leu Val Ile Gln Gly Pro Leu Lys Pro Ala Asp 140 145 150	966		
40	gtt acg gtt gat ggg tcc tta agc tct cag ttc ctt aca ggt ttg ttg Val Thr Val Asp Gly Ser Leu Ser Ser Gln Phe Leu Thr Gly Leu Leu 155 160 165	1014		
45	ctt gca tat gcg gcc gca gat gca agc gat gtt gcg ata aaa gta acg Leu Ala Tyr Ala Ala Ala Asp Ala Ser Asp Val Ala Ile Lys Val Thr 170 175 180 185	1062		
50	aat ctc aaa agc cgt ccg tat atc gat ctt aca ctg gat gtg atg aag Asn Leu Lys Ser Arg Pro Tyr Ile Asp Leu Thr Leu Asp Val Met Lys 190 195 200	1110		
55	cgg ttt ggt ttg aag act ccc gag aat cga aac tat gaa gag ttt tat Arg Phe Gly Leu Lys Thr Pro Glu Asn Arg Asn Tyr Glu Glu Phe Tyr 205 210 215	1158		
60	ttc aaa gcc ggg aat gta tat gat gaa acg aaa atg caa cga tac acc Phe Lys Ala Gly Asn Val Tyr Asp Glu Thr Lys Met Gln Arg Tyr Thr 220 225 230	1206		
65	gta gaa ggc gac tgg agc ggt ggt gct ttt tta ctg gta gcg ggg gct Val Glu Gly Asp Trp Ser Gly Gly Ala Phe Leu Leu Val Ala Gly Ala 235 240 245	1254		
70	att gcc ggg ccg atc acg gta aga ggt ttg gat ata gct tcg acg cag Ile Ala Gly Pro Ile Thr Val Arg Gly Leu Asp Ile Ala Ser Thr Gln 250 255 260 265	1302		
75	gct gat aaa gcg atc gtt cag gct ttg atg agt gcg aac gca ggt att	1350		

EP 1 840 210 A1

	Ala Asp Lys Ala Ile Val Gln Ala Leu Met Ser Ala Asn Ala Gly Ile	
	270 275 280	
5	gcg att gat gca aaa gag atc aaa ctt cat cct gct gat ctc aat gca Ala Ile Asp Ala Lys Glu Ile Lys Leu His Pro Ala Asp Leu Asn Ala	1398
	285 290 295	
10	ttt gaa ttt gat gct act gat tgc ccg gat ctt ttt ccg cca ttg gtt Phe Glu Phe Asp Ala Thr Asp Cys Pro Asp Leu Phe Pro Pro Leu Val	1446
	300 305 310	
15	gct ttg gcg tct tat tgc aaa gga gaa aca aag atc aaa ggc gta agc Ala Leu Ala Ser Tyr Cys Lys Gly Glu Thr Lys Ile Lys Gly Val Ser	1494
	315 320 325	
20	agg ctg gcg cat aaa gaa agt gac aga gga ttg acg ctg cag gac gag Arg Leu Ala His Lys Glu Ser Asp Arg Gly Leu Thr Leu Gln Asp Glu	1542
	330 335 340 345	
25	ttc ggg aaa atg ggt gtt gaa atc cac ctt gag gga gat ctg atg cgc Phe Gly Lys Met Gly Val Glu Ile His Leu Glu Gly Asp Leu Met Arg	1590
	350 355 360	
30	gtg atc gga ggg aaa ggc gta aaa gga gct gaa gtt agt tca agg cac Val Ile Gly Gly Lys Gly Val Lys Gly Ala Glu Val Ser Ser Arg His	1638
	365 370 375	
35	gat cat cgc att gcg atg gct tgc gcg gtg gct gct tta aaa gct gtg Asp His Arg Ile Ala Met Ala Cys Ala Val Ala Ala Leu Lys Ala Val	1686
	380 385 390	
40	ggt gaa aca acc atc gaa cat gca gaa gcg gtg aat aaa tcc tac ccg Gly Glu Thr Thr Ile Glu His Ala Glu Ala Val Asn Lys Ser Tyr Pro	1734
	395 400 405	
45	gat ttt tac agc gat ctt aaa caa ctt ggc ggt gtt gta tct tta aac Asp Phe Tyr Ser Asp Leu Lys Gln Leu Gly Gly Val Val Ser Leu Asn	1782
	410 415 420 425	
50	cat caa ttt aat ttc tca tgaatagctt cggccgcac ttcagggtgc His Gln Phe Asn Phe Ser	1830
	430	
55	atatTTTTGG cgaatcacat ggtgaatcag taggcacgt tattgatggt tgcctgctg	1890
	gtctgtcatt gtccgaagaa gatc	1914
	<210> 3 <211> 431 <212> PRT <213> Pseudomonas putida P.P4G-1	
	<400> 3 Met Lys Val Thr Ile Gln Pro Gly Asp Leu Thr Gly Ile Ile Gln Ser	
	1 5 10 15	
	Pro Ala Ser Lys Ser Ser Met Gln Arg Ala Cys Ala Ala Ala Leu Val	

EP 1 840 210 A1

	20	25	30
5	Ala Lys Gly Ile Ser Glu Ile	Ile Asn Pro Gly His	Ser Asn Asp Asp
	35	40	45
10	Lys Ala Ala Arg Asp Ile Val Ser Arg Leu Gly Ala Arg Leu Glu Asp		
	50	55	60
15	Gln Pro Asp Gly Ser Leu Gln Ile Thr Ser Glu Gly Val Lys Pro Val		
	65	70	75 80
20	Ala Pro Phe Ile Asp Cys Gly Glu Ser Gly Leu Ser Ile Arg Met Phe		
	85	90	95
25	Thr Pro Ile Val Ala Leu Ser Lys Glu Glu Val Thr Ile Lys Gly Ser		
	100	105	110
30	Gly Ser Leu Val Thr Arg Pro Met Asp Phe Phe Asp Glu Ile Leu Pro		
	115	120	125
35	His Leu Gly Val Lys Val Lys Ser Asn Gln Gly Lys Leu Pro Leu Val		
	130	135	140
40	Ile Gln Gly Pro Leu Lys Pro Ala Asp Val Thr Val Asp Gly Ser Leu		
	145	150	155 160
45	Ser Ser Gln Phe Leu Thr Gly Leu Leu Leu Ala Tyr Ala Ala Ala Asp		
	165	170	175
50	Ala Ser Asp Val Ala Ile Lys Val Thr Asn Leu Lys Ser Arg Pro Tyr		
	180	185	190
55	Ile Asp Leu Thr Leu Asp Val Met Lys Arg Phe Gly Leu Lys Thr Pro		
	195	200	205
60	Glu Asn Arg Asn Tyr Glu Glu Phe Tyr Phe Lys Ala Gly Asn Val Tyr		
	210	215	220
65	Asp Glu Thr Lys Met Gln Arg Tyr Thr Val Glu Gly Asp Trp Ser Gly		
	225	230	235 240
70	Gly Ala Phe Leu Leu Val Ala Gly Ala Ile Ala Gly Pro Ile Thr Val		
	245	250	255

EP 1 840 210 A1

Arg Gly Leu Asp Ile Ala Ser Thr Gln Ala Asp Lys Ala Ile Val Gln
 260 265 270
 5
 Ala Leu Met Ser Ala Asn Ala Gly Ile Ala Ile Asp Ala Lys Glu Ile
 275 280 285
 10
 Lys Leu His Pro Ala Asp Leu Asn Ala Phe Glu Phe Asp Ala Thr Asp
 290 295 300
 15
 Cys Pro Asp Leu Phe Pro Pro Leu Val Ala Leu Ala Ser Tyr Cys Lys
 305 310 315 320
 Gly Glu Thr Lys Ile Lys Gly Val Ser Arg Leu Ala His Lys Glu Ser
 325 330 335
 20
 Asp Arg Gly Leu Thr Leu Gln Asp Glu Phe Gly Lys Met Gly Val Glu
 340 345 350
 25
 Ile His Leu Glu Gly Asp Leu Met Arg Val Ile Gly Gly Lys Gly Val
 355 360 365
 30
 Lys Gly Ala Glu Val Ser Ser Arg His Asp His Arg Ile Ala Met Ala
 370 375 380
 35
 Cys Ala Val Ala Ala Leu Lys Ala Val Gly Glu Thr Thr Ile Glu His
 385 390 395 400
 40
 Ala Glu Ala Val Asn Lys Ser Tyr Pro Asp Phe Tyr Ser Asp Leu Lys
 405 410 415
 45
 Gln Leu Gly Gly Val Val Ser Leu Asn His Gln Phe Asn Phe Ser
 420 425 430
 <210> 4
 <211> 20
 <212> DNA
 <213> artificial sequence
 <400> 4
 agagtttgat catggctcag 20
 50
 <210> 5
 <211> 22
 <212> DNA
 <213> artificial sequence
 55
 <400> 5

EP 1 840 210 A1

	tacggttacc ttgttacgac tt	22
5	<210> 6 <211> 34 <212> DNA <213> artificial sequence	
10	<400> 6 cgggatccta agtaagtga agtaacaata cagc	34
15	<210> 7 <211> 29 <212> DNA <213> artificial sequence	
20	<400> 7 cgggatccct tcttcggaca atgacagac	29
25	<210> 8 <211> 33 <212> DNA <213> artificial sequence	
30	<400> 8 cgggatccgt taatgccgaa attttgctta atc	33
35	<210> 9 <211> 30 <212> DNA <213> artificial sequence	
40	<400> 9 cgggatccag gtccgaaaaa aaacgccgac	30
45	<210> 10 <211> 1436 <212> DNA <213> Escherichia coli	
50	<400> 10 gttaatgccg aaatTTTgct taatccccac agccagcctg tggggTTTT atttctgttg	60
55	tagagagttg agttcatgga atccctgacg ttacaacca tcgctcgtgt cgatggcact	120
	attaatctgc ccggttccaa gagcgTTTct aaccgcgctt tattgctggc ggcattagca	180
	cacggcaaaa cagtattaac caatctgctg gatagcgatg acgtgcgcca tatgctgaat	240
	gcattaacag cgtaggggt aagctatacg ctttcagccg atcgtagcg ttgcgaaatt	300
	atcggtaacg gcggtccatt acacgcagaa ggtgccctgg agttgttctt cggtaacgcc	360
	ggaacggcaa tgcgtccgct ggccggcagct ctttgtctgg gtagcaatga tattgtgctg	420

EP 1 840 210 A1

	accggtgagc cgctatgaa agaacgcccg attggtcacg tggatgatgc gctgcgcctg	480
	ggcggggcga agatcactta cctggaacaa gaaaattatc cgccgttgcg ttacagggc	540
5	ggctttactg gcggcaacgt tgacgttgat ggctccgttt ccagccaatt cctcaccgca	600
	ctgttaatga ctgcgcctct tgcgccgga gatacggatg ttcgtattaa aggcgatctg	660
10	gtttctaaac cttatatcga catcacactc aatctgatga agacgtttgg tgttgaatt	720
	gaaaatcagc actatcaaca atttgtcgta aaaggcgggc agtcttatca gtctccgggt	780
	acttatttgg tcgaaggcga tgcattctcg gcttcttact ttctggcagc agcagcaatc	840
15	aaaggcggca ctgtaaaagt gaccggtatt ggacgtaaca gtatgcaggg tgatattcgc	900
	tttctgatg tgctggaaaa aatgggcgcg accatttgcg gggcgatga ttatatttcc	960
	tgcacgcgtg gtgaactgaa cgctattgat atggatatga accatattcc tgatgcggcg	1020
20	atgaccattg ccacggcggc gttatttga aaaggcacca ccacgtcgcg caatatctat	1080
	aactggcgtg ttaaagagac cgatcgctg tttgcgatgg caacagaact gcgtaaagtc	1140
25	ggcgcggaag tggaagaggg gcacgattac attcgtatca ctctccgga aaaactgaac	1200
	tttccgaga tcgcgacata caatgatcac cggatggcga tgtgtttctc gctggtggcg	1260
	ttgtcagata caccagtgc gattcttgat cccaatgca cggccaaaac atttccggat	1320
30	tatttcgagc agctggcgcg gattagccag gcagcctgaa tgaacaacgg gcaataaata	1380
	gccaaatctt tctttatcaa aacgtcggca cattgtcggc gtttttttc ggacct	1436
35	<210> 11 <211> 30 <212> DNA <213> artificial sequence	
40	<220> <223> upstream primer for Escherichia coli aroA	
	<400> 11 cgggatccag gtccgaaaaa aaacgccgac	30
45	<210> 12 <211> 28 <212> DNA <213> artificial sequence	
50	<220> <223> downstream primer for Escherichia coli aroA	
	<400> 1 2 cgggatccat ggaatccctg acgttaca	28
55	<210> 13 <211> 29	

EP 1 840 210 A1

	<212> DNA	
	<213> artificial sequence	
5	<220>	
	<223> pACYC series upstream primers for Escherichia coli aroA	
	<400> 13	
	tgagtgactg actttaagaa ggagatata	29
10	<210> 14	
	<211> 28	
	<212> DNA	
	<213> artificial sequence	
15	<220>	
	<223> pBR322 series downstream primers for Escherichia coli aroA	
	<400> 14	
	gccacgatgc gtccggcgta gaggatcc	28
20	<210> 15	
	<211> 31	
	<212> DNA	
	<213> artificial sequence	
25	<220>	
	<223> N218 downstream primers for Escherichia coli aroA	
	<400> 15	
	cgggatcctc actgattttc aatttcaaca c	31
30	<210> 16	
	<211> 31	
	<212> DNA	
	<213> artificial sequence	
35	<220>	
	<223> C219 downstream primers for Escherichia coli aroA	
	<400> 16	
	catgccatgg gtcaacaatt tgcgtaaaa g	31
40	<210> 17	
	<211> 32	
	<212> DNA	
	<213> artificial sequence	
45	<220>	
	<223> N227 downstream primers for Escherichia coli aroA	
	<400> 17	
	cgggatcctc agccttttac gacaaattgt tg	32
50	<210> 18	
	<211> 31	
	<212> DNA	
	<213> artificial sequence	
55	<220>	

EP 1 840 210 A1

	<223> N234 downstream primers for Escherichia coli aroA	
5	<400> 18 cgggatacctc agacaaattg ttgatagtgc t	31
	<210> 19 <211> 31 <212> DNA <213> artificial sequence	
10	<220> <223> N245 downstream primers for Escherichia coli aroA	
15	<400> 19 cgggatacctc acagcaataa agcgcggtta g	31
	<210> 20 <211> 29 <212> DNA <213> artificial sequence	
20	<220> <223> C228 upstream primers for Escherichia coli aroA	
25	<400> 20 catgccatgg ggcagtctta tcagtctcc	29
	<210> 21 <211> 28 <212> DNA <213> artificial sequence	
30	<220> <223> C235 upstream primers for Escherichia coli aroA	
35	<400> 21 catgccatgg gtacttattt ggtcgaag	28
	<210> 22 <211> 32 <212> DNA <213> artificial sequence	
40	<220> <223> N245 downstream primers for Escherichia coli aroA	
45	<400> 22 cgggatacctc acgaagatgc atcgcttcg ac	32
	<210> 23 <211> 31 <212> DNA <213> artificial sequence	
50	<220> <223> N240 downstream primers for Escherichia coli aroA	
55	<400> 23 cgggatacctc attcgaccaa ataagtaccg g	31

EP 1 840 210 A1

5
 <210> 24
 <211> 29
 <212> DNA
 <213> artificial sequence

10
 <220>
 <223> C241 upstream primers for Escherichia coli aroA
 <400> 24
 catgcatgg gcgatgcatc ttcggcttc 29

15
 <210> 25
 <211> 31
 <212> DNA
 <213> artificial sequence

20
 <220>
 <223> pKU2159 downstream primer
 <400> 25
 gttactcgag ggctgcctgg ctaatccgcg c 31

25
 <210> 26
 <211> 23
 <212> DNA
 <213> artificial sequence

30
 <220>
 <223> pKU2223 downstream primer
 <400> 26
 ggaattccat gaactcaact ctc 23

35
 <210> 27
 <211> 25
 <212> DNA
 <213> artificial sequence

40
 <220>
 <223> pKU2223 upstream primer
 <400> 27
 cgggatcctc aacgataacg gctcc 25

45
 <210> 28
 <211> 23
 <212> DNA
 <213> artificial sequence

50
 <220>
 <223> pKU2224 upstream primer
 <400> 28
 ggaattcaaa tatgtatccg ctc 23

55
 <210> 29
 <211> 26
 <212> DNA

EP 1 840 210 A1

	<213> artificial sequence	
	<220>	
5	<223> pKU2224 downstream primer	
	<400> 29	
	cccaagcttg gtctgacagt taccaa	26
10	<210> 30	
	<211> 25	
	<212> DNA	
	<213> artificial sequence	
	<220>	
15	<223> pKU2225 upstream primer	
	<400> 30	
	cccaagcttg atcccaaagt cacgg	25
20	<210> 31	
	<211> 27	
	<212> DNA	
	<213> artificial sequence	
25	<220>	
	<223> pKU2225 downstream primer	
	<400> 31	
30	acgaagtcga cgtggcccag ttcattg	27
	<210> 32	
	<211> 32	
	<212> DNA	
	<213> artificial sequence	
35	<220>	
	<223> N67 downstream primers for Escherichia coli aroA	
	<400> 32	
40	cttaggatcc tcatgaaagc gtatagctta cc	32
	<210> 33	
	<211> 32	
	<212> DNA	
	<213> artificial sequence	
45	<220>	
	<223> N73 downstream primers for Escherichia coli aroA	
	<400> 33	
50	gtcaggatcc tcagcaacgc gtacgatcgg ct	32
	<210> 34	
	<211> 32	
	<212> DNA	
	<213> artificial sequence	
55	<220>	

EP 1 840 210 A1

	<223> N85 downstream primers for Escherichia coli aroA	
5	<400> 34 gtcaggatcc tcattctgcg tgtaatggac cg	32
	<210> 35 <211> 33 <212> DNA <213> artificial sequence	
10	<220> <223> N104 downstream primers for Escherichia coli aroA	
15	<400> 35 cttaggatcc tcatgccgcc agcggacgca ttg	33
	<210> 36 <211> 33 <212> DNA <213> artificial sequence	
20	<220> <223> N154 downstream primers for Escherichia coli aroA	
25	<400> 36 cgcaggatcc tcaactgtaaa cgcaacggcg gat	33
	<210> 37 <211> 31 <212> DNA <213> artificial sequence	
30	<220> <223> N165 downstream primers for Escherichia coli aroA	
35	<400> 37 cgggatcctc aatcaacgtc aacgttgccg c	31
	<210> 38 <211> 35 <212> DNA <213> artificial sequence	
40	<220> <223> N182 downstream primers for Escherichia coli aroA	
45	<400> 38 ggcaggatcc tcaaagaggc gcagtcatta acagt	35
	<210> 39 <211> 32 <212> DNA <213> artificial sequence	
50	<220> <223> N184 downstream primers for Escherichia coli aroA	
55	<400> 39 tccaggatcc tcacggcgca agaggcgag tc	32

EP 1 840 210 A1

5
 <210> 40
 <211> 32
 <212> DNA
 <213> artificial sequence

10
 <220>
 <223> N224 downstream primers for Escherichia coli aroA
 <400> 40
 cgggatacctc agacaaattg ttgatagtgc tg 32

15
 <210> 41
 <211> 31
 <212> DNA
 <213> artificial sequence

20
 <220>
 <223> N238 downstream primers for Escherichia coli aroA
 <400> 41
 cgggatacctc aagtacccgg agactgataa g 31

25
 <210> 42
 <211> 31
 <212> DNA
 <213> artificial sequence

30
 <220>
 <223> N259 downstream primer of E.coli aroA
 <400> 42
 cgggatacctc aagtgccgcc ttgattgct g 31

35
 <210> 43
 <211> 32
 <212> DNA
 <213> artificial sequence

40
 <220>
 <223> N298 downstream primers for Escherichia coli aroA
 <400> 43
 cgggatacctc aacgcgtgca ggaaatataa tc 32

45
 <210> 44
 <211> 31
 <212> DNA
 <213> artificial sequence

50
 <220>
 <223> N371 downstream primers for Escherichia coli aroA
 <400> 44
 cgggatacctc acggaggagt gatacgaatg t 31

55
 <210> 45
 <211> 35
 <212> DNA

EP 1 840 210 A1

	<213> artificial sequence	
	<220>	
5	<223> N376 downstream primers for Escherichia coli aroA	
	<400> 45	
	ccgtggatcc tcaaaagttc agtttttccg gagga	35
10	<210> 46	
	<211> 27	
	<212> DNA	
	<213> artificial sequence	
	<220>	
15	<223> C31 upstream primers for Escherichia coli aroA	
	<400> 46	
	catgccatgg cggcattagc acacggc	27
20	<210> 47	
	<211> 29	
	<212> DNA	
	<213> artificial sequence	
	<220>	
25	<223> C68 upstream primers for Escherichia coli aroA	
	<400> 47	
	tatgccatgg ccgatcgtac gcgttgcca	29
30	<210> 48	
	<211> 31	
	<212> DNA	
	<213> artificial sequence	
	<220>	
35	<223> C74 upstream primers for Escherichia coli aroA	
	<400> 48	
	tgcacccatgg aaattatcgg taacggcggc c	31
40	<210> 49	
	<211> 31	
	<212> DNA	
	<213> artificial sequence	
	<220>	
45	<223> C86 upstream primers for Escherichia coli aroA	
	<400> 49	
	tattccatgg gtgccctgga gttgttcctc g	31
50	<210> 50	
	<211> 28	
	<212> DNA	
	<213> artificial sequence	
	<220>	
55	<223> C105 upstream primers for Escherichia coli aroA	

EP 1 840 210 A1

5	<400> 50 cggaccatgg ctctttgtct gggtagca	28
	<210> 51 <211> 31 <212> DNA <213> artificial sequence	
10	<220> <223> C155 upstream primers for Escherichia coli aroA	
15	<400> 51 ttatccatgg gcggctttac tggcggcaac g	31
	<210> 52 <211> 30 <212> DNA <213> artificial sequence	
20	<220> <223> C166 upstream primers for Escherichia coli aroA	
25	<400> 52 catgccatgg gctccgtttc cagccaattc	30
	<210> 53 <211> 29 <212> DNA <213> artificial sequence	
30	<220> <223> C183 upstream primers for Escherichia coli aroA	
35	<400> 53 tgagccatgg cgccggaaga tacggtgat	29
	<210> 54 <211> 30 <212> DNA <213> artificial sequence	
40	<220> <223> C185 upstream primers for Escherichia coli aroA	
45	<400> 54 gcgtccatgg aagatacggg gattcgtatt	30
	<210> 55 <211> 29 <212> DNA <213> artificial sequence	
50	<220> <223> C225 upstream primers for Escherichia coli aroA	
55	<400> 55 catgccatgg taaaaggcgg gcagtcctta	29

EP 1 840 210 A1

5	<210> 56 <211> 26 <212> DNA <213> artificial sequence	
	<220> <223> C239 upstream primers for Escherichia coli aroA	
10	<400> 56 catgccatgg tcgaaggcga tgcac	26
	<210> 57 <211> 28 <212> DNA <213> artificial sequence	
15	<220> <223> C246 upstream primers for Escherichia coli aroA	
20	<400> 57 catgccatgg cttcttactt tctggcag	28
	<210> 58 <211> 29 <212> DNA <213> artificial sequence	
25	<220> <223> C260 upstream primers for Escherichia coli aroA	
30	<400> 58 catgccatgg taaaagtac cggtattgg	29
	<210> 59 <211> 29 <212> DNA <213> artificial sequence	
35	<220> <223> C299 upstream primers for Escherichia coli aroA	
40	<400> 59 catgccatgg gtgaactgaa cgctattga	29
	<210> 60 <211> 31 <212> DNA <213> artificial sequence	
45	<220> <223> primer	
50	<400> 60 tgagtgactg aaagtgaag taacaataga g	31
	<210> 61 <211> 29 <212> DNA <213> artificial sequence	
55		

	<220>		
	<223>	primer	
5			
	<400>	61	
		cgggatccct tcttcggaca atgacagac	29
	<210>	62	
10	<211>	31	
	<212>	DNA	
	<213>	artificial sequence	
	<220>		
15	<223>	primer ppN2083' (BamHI) N208.C209+	
	<400>	62	
		cgggatcctc agggagtctt caaaccaaac c	31
	<210>	63	
20	<211>	28	
	<212>	DNA	
	<213>	artificial sequence	
	<220>		
25	<223>	primer ppC2095' (NcoI) N208/C209 +	
	<400>	63	
		catgccatgg agaatcgaaa ctatgaag	28
30	<210>	64	
	<211>	31	
	<212>	DNA	
	<213>	artificial sequence	
	<220>		
35	<223>	primer ppN2143' (BamHI) N214/C215 +	
	<400>	64	
		cgggatcctc attcatagtt tcgattctcg g	31
40	<210>	65	
	<211>	29	
	<212>	DNA	
	<213>	artificial sequence	
45	<220>		
	<223>	primer ppC2155' (NcoI) N214/C215+	
	<400>	65	
		catgccatgg agttttatattt caaagccgg	29
50	<210>	66	
	<211>	33	
	<212>	DNA	
	<213>	artificial sequence	
55	<220>		
	<223>	primer ppN2193' (BamHI) N219/C220 ++	

	<400> 66 cgggatacctc atttgaaata aaactcttca tag	33
5	<210> 67 <211> 32 <212> DNA <213> artificial sequence	
10	<220> <223> primer ppC2205' (NcoI) N219/C220 ++	
	<400> 67 catgccatgg cggggaatgt atatgatgaa ac	32
15	<210> 68 <211> 33 <212> DNA <213> artificial sequence	
20	<220> <223> primer ppN2223' (BamHI) N222/C223 +	
	<400> 68 cgggatacctc aattcccggc ttgaaataa aac	33
25	<210> 69 <211> 30 <212> DNA <213> artificial sequence	
30	<220> <223> primer ppC2235' (NcoI) N222/C223 +	
	<400> 69 catgccatgg tatatgatga aacgaaaatg	30
35	<210> 70 <211> 30 <212> DNA <213> artificial sequence	
40	<220> <223> primer ppN2243' (BamHI) N224/C225 ++	
	<400> 70 cgggatacctc aatatacatt cccggctttg	30
45	<210> 71 <211> 29 <212> DNA <213> artificial sequence	
50	<220> <223> primer ppC2255' (NcoI) N224/C225 ++	
	<400> 71 catgccatgg atgaaacgaa aatgcaacg	29
55	<210> 72	

	<211> 31	
	<212> DNA	
	<213> artificial sequence	
5	<220>	
	<223> primer ppN2333' (BamHI) N233/C234	
	<400> 72	
10	cgggatacctc aggtgtatcg ttgcattttc g	31
	<210> 73	
	<211> 28	
	<212> DNA	
15	<213> artificial sequence	
	<220>	
	<223> primer ppC2345' (NcoI) N233/C234	
	<400> 73	
20	catgccatgg tagaaggcga ctggagcg	28
	<210> 74	
	<211> 33	
	<212> DNA	
25	<213> artificial sequence	
	<220>	
	<223> primer ppN2343' (BamHI) N234/C235	
	<400> 74	
30	cgggatacctc atacggtgta tcgttgcat ttc	33
	<210> 75	
	<211> 29	
	<212> DNA	
35	<213> artificial sequence	
	<220>	
	<223> primer ppC2355' (NcoI) N234/C235	
	<400> 75	
40	catgccatgg aaggcgactg gagcgggtgg	29
	<210> 76	
	<211> 32	
	<212> DNA	
45	<213> artificial sequence	
	<220>	
	<223> primer ppN2363' (BamHI) N236/C237	
	<400> 76	
50	cgggatacctc agccttctac ggtgtatcgt tg	32
	<210> 77	
	<211> 29	
	<212> DNA	
55	<213> artificial sequence	

EP 1 840 210 A1

<220>
<223> primer ppc2375' (NcoI) N236/C237

<400> 77
catgcatgg actggagcgg tggcgcttt 29

<210> 78
<211> 1284
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(1284)
<223>

<400> 78
atg gaa tcc ctg acg tta caa ccc atc gct cgt gtc gat ggc act att 48
Met Glu Ser Leu Thr Leu Gln Pro Ile Ala Arg Val Asp Gly Thr Ile
1 5 10 15

aat ctg ccc ggt tcc aag agc gtt tct aac cgc gct tta ttg ctg gcg 96
Asn Leu Pro Gly Ser Lys Ser Val Ser Asn Arg Ala Leu Leu Leu Ala
20 25 30

gca tta gca cac ggc aaa aca gta tta acc aat ctg ctg gat agc gat 144
Ala Leu Ala His Gly Lys Thr Val Leu Thr Asn Leu Leu Asp Ser Asp
35 40 45

gac gtg cgc cat atg ctg aat gca tta aca gcg tta ggg gta agc tat 192
Asp Val Arg His Met Leu Asn Ala Leu Thr Ala Leu Gly Val Ser Tyr
50 55 60

acg ctt tca gcc gat cgt acg cgt tgc gaa att atc ggt aac ggc ggt 240
Thr Leu Ser Ala Asp Arg Thr Arg Cys Glu Ile Ile Gly Asn Gly Gly
65 70 75 80

cca tta cac gca gaa ggt gcc ctg gag ttg ttc ctc ggt aac gcc gga 288
Pro Leu His Ala Glu Gly Ala Leu Glu Leu Phe Leu Gly Asn Ala Gly
85 90 95

acg gca atg cgt ccg ctg gcg gca gct ctt tgt ctg ggt agc aat gat 336
Thr Ala Met Arg Pro Leu Ala Ala Ala Leu Cys Leu Gly Ser Asn Asp
100 105 110

att gtg ctg acc ggt gag ccg cgt atg aaa gaa cgc ccg att ggt cat 384
Ile Val Leu Thr Gly Glu Pro Arg Met Lys Glu Arg Pro Ile Gly His
115 120 125

ctg gtg gat gcg ctg cgc ctg ggc ggg gcg aag atc act tac ctg gaa 432
Leu Val Asp Ala Leu Arg Leu Gly Gly Ala Lys Ile Thr Tyr Leu Glu
130 135 140

caa gaa aat tat ccg ccg ttg cgt tta cag ggc ggc ttt act ggc ggc 480
Gln Glu Asn Tyr Pro Pro Leu Arg Leu Gln Gly Gly Phe Thr Gly Gly
145 150 155 160

EP 1 840 210 A1

	aac gtt gac gtt gat ggc tcc gtt tcc agc caa ttc ctc acc gca ctg	528
	Asn Val Asp Val Asp Gly Ser Val Ser Ser Gln Phe Leu Thr Ala Leu	
	165 170 175	
5	tta atg act gcg cct ctt gcg ccg gaa gat acg gtg att cgt att aaa	576
	Leu Met Thr Ala Pro Leu Ala Pro Glu Asp Thr Val Ile Arg Ile Lys	
	180 185 190	
10	ggc gat ctg gtt tct aaa cct tat atc gac atc aca ctc aat ctg atg	624
	Gly Asp Leu Val Ser Lys Pro Tyr Ile Asp Ile Thr Leu Asn Leu Met	
	195 200 205	
15	aag acg ttt ggt gtt gaa att gaa aat cag cac tat caa caa ttt gtc	672
	Lys Thr Phe Gly Val Glu Ile Glu Asn Gln His Tyr Gln Gln Phe Val	
	210 215 220	
20	gta aaa ggc ggc cag tct tat cag tct ccg ggt act tat ttg gtc gaa	720
	Val Lys Gly Gly Gln Ser Tyr Gln Ser Pro Gly Thr Tyr Leu Val Glu	
	225 230 235 240	
25	ggc gat gca tct tcg gct tct tac ttt ctg gca gca gca gca atc aaa	768
	Gly Asp Ala Ser Ser Ala Ser Tyr Phe Leu Ala Ala Ala Ala Ile Lys	
	245 250 255	
30	ggc ggc act gta aaa gtg acc ggt att gga cgt aac agt atg cag ggt	816
	Gly Gly Thr Val Lys Val Thr Gly Ile Gly Arg Asn Ser Met Gln Gly	
	260 265 270	
35	gat att cgc ttt gct gat gtg ctg gaa aaa atg ggc gcg acc att tgc	864
	Asp Ile Arg Phe Ala Asp Val Leu Glu Lys Met Gly Ala Thr Ile Cys	
	275 280 285	
40	tgg ggc gat gat tat att tcc tgc acg cgt ggt gaa ctg aac gct att	912
	Trp Gly Asp Asp Tyr Ile Ser Cys Thr Arg Gly Glu Leu Asn Ala Ile	
	290 295 300	
45	gat atg gat atg aac cat att cct gat gcg gcg atg acc att gcc acg	960
	Asp Met Asp Met Asn His Ile Pro Asp Ala Ala Met Thr Ile Ala Thr	
	305 310 315 320	
50	gcg gcg tta ttt gca aaa ggc acc acc acg ctg cgc aat atc tat aac	1008
	Ala Ala Leu Phe Ala Lys Gly Thr Thr Thr Leu Arg Asn Ile Tyr Asn	
	325 330 335	
55	tgg cgt gtt aaa gag acc gat cgc ctg ttt gcg atg gca aca gaa ctg	1056
	Trp Arg Val Lys Glu Thr Asp Arg Leu Phe Ala Met Ala Thr Glu Leu	
	340 345 350	
60	cgt aaa gtc ggc gcg gaa gtg gaa gag ggg cac gat tac att cgt atc	1104
	Arg Lys Val Gly Ala Glu Val Glu Glu Gly His Asp Tyr Ile Arg Ile	
	355 360 365	
65	act cct ccg gaa aaa ctg aac ttt gcc gag atc gcg aca tac aat gat	1152
	Thr Pro Pro Glu Lys Leu Asn Phe Ala Glu Ile Ala Thr Tyr Asn Asp	
	370 375 380	
70	cac cgg atg gcg atg tgt ttc tcg ctg gtg gcg ttg tca gat aca cca	1200
	His Arg Met Ala Met Cys Phe Ser Leu Val Ala Leu Ser Asp Thr Pro	
	385 390 395 400	

EP 1 840 210 A1

gtg acg att ctt gat ccc aaa tgc acg gcc aaa aca ttt ccg gat tat 1248
Val Thr Ile Leu Asp Pro Lys Cys Thr Ala Lys Thr Phe Pro Asp Tyr
5 405 410 415

ttc gag cag ctg gcg cgg att agc cag gca gcc tga 1284
Phe Glu Gln Leu Ala Arg Ile Ser Gln Ala Ala
420 425

10

<210> 79
<211> 427
<212> PRT
<213> Escherichia coli

15

<400> 79
Met Glu Ser Leu Thr Leu Gln Pro Ile Ala Arg Val Asp Gly Thr Ile
1 5 10 15

20

Asn Leu Pro Gly Ser Lys Ser Val Ser Asn Arg Ala Leu Leu Leu Ala
20 25 30

Ala Leu Ala His Gly Lys Thr Val Leu Thr Asn Leu Leu Asp Ser Asp
35 40 45

25

Asp Val Arg His Met Leu Asn Ala Leu Thr Ala Leu Gly Val Ser Tyr
50 55 60

Thr Leu Ser Ala Asp Arg Thr Arg Cys Glu Ile Ile Gly Asn Gly Gly
65 70 75 80

30

Pro Leu His Ala Glu Gly Ala Leu Glu Leu Phe Leu Gly Asn Ala Gly
85 90 95

Thr Ala Met Arg Pro Leu Ala Ala Ala Leu Cys Leu Gly Ser Asn Asp
100 105 110

35

Ile Val Leu Thr Gly Glu Pro Arg Met Lys Glu Arg Pro Ile Gly His
115 120 125

Leu Val Asp Ala Leu Arg Leu Gly Gly Ala Lys Ile Thr Tyr Leu Glu
130 135 140

40

Gln Glu Asn Tyr Pro Pro Leu Arg Leu Gln Gly Gly Phe Thr Gly Gly
145 150 155 160

Asn Val Asp Val Asp Gly Ser Val Ser Ser Gln Phe Leu Thr Ala Leu
165 170 175

45

Leu Met Thr Ala Pro Leu Ala Pro Glu Asp Thr Val Ile Arg Ile Lys
180 185 190

Gly Asp Leu Val Ser Lys Pro Tyr Ile Asp Ile Thr Leu Asn Leu Met
195 200 205

50

Lys Thr Phe Gly Val Glu Ile Glu Asn Gln His Tyr Gln Gln Phe Val
210 215 220

55

Val Lys Gly Gly Gln Ser Tyr Gln Ser Pro Gly Thr Tyr Leu Val Glu

	225				230					235					240
5	Gly	Asp	Ala	Ser	Ser	Ala	Ser	Tyr	Phe	Leu	Ala	Ala	Ala	Ile	Lys
					245					250				255	
	Gly	Gly	Thr	Val	Lys	Val	Thr	Gly	Ile	Gly	Arg	Asn	Ser	Met	Gln
				260					265					270	
10	Asp	Ile	Arg	Phe	Ala	Asp	Val	Leu	Glu	Lys	Met	Gly	Ala	Thr	Ile
			275					280					285		Cys
	Trp	Gly	Asp	Asp	Tyr	Ile	Ser	Cys	Thr	Arg	Gly	Glu	Leu	Asn	Ala
15		290					295					300			Ile
	Asp	Met	Asp	Met	Asn	His	Ile	Pro	Asp	Ala	Ala	Met	Thr	Ile	Ala
	305					310					315				320
20	Ala	Ala	Leu	Phe	Ala	Lys	Gly	Thr	Thr	Thr	Leu	Arg	Asn	Ile	Tyr
					325						330				335
	Trp	Arg	Val	Lys	Glu	Thr	Asp	Arg	Leu	Phe	Ala	Met	Ala	Thr	Glu
				340					345					350	Leu
25	Arg	Lys	Val	Gly	Ala	Glu	Val	Glu	Glu	Gly	His	Asp	Tyr	Ile	Arg
			355					360					365		Ile
	Thr	Pro	Pro	Glu	Lys	Leu	Asn	Phe	Ala	Glu	Ile	Ala	Thr	Tyr	Asn
30		370					375					380			Asp
	His	Arg	Met	Ala	Met	Cys	Phe	Ser	Leu	Val	Ala	Leu	Ser	Asp	Thr
	385					390					395				400
35	Val	Thr	Ile	Leu	Asp	Pro	Lys	Cys	Thr	Ala	Lys	Thr	Phe	Pro	Asp
					405					410					415
	Phe	Glu	Gln	Leu	Ala	Arg	Ile	Ser	Gln	Ala	Ala				
40				420					425						

Claims

- 45 1. A 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) protein fragment selected from a pair of EPSPS protein fragments, wherein the fragment pair is comprised of two fragments which are capable of joining into full length EPSPS and reconstituting EPSPS activity by complementation without help of any joint structure.
- 50 2. The fragment according to claim 1, wherein the fragment pair is derived from division of EPSPS at a position selected from the group consisting of: a position in a joining region between folding units, a position in a joining region between an α -helix and a β -sheet, a position between two β -sheets, a position in a β -sheet or a α -helix, preferably a position in a joining region between folding units.
- 55 3. The fragment according to claims 1 or 2, wherein the EPSPS is a wild type EPSPS or an active variant thereof in which one or more amino acid residue is added, deleted, and/or substituted, preferably the EPSPS is an *E. coli* wild type EPSPS or a glyphosate-tolerant active variant thereof, or preferably the EPSPS is a *P. putida* glyphosate-tolerant EPSPS.

4. The fragment according to claim 1, wherein the fragment pair is selected from a group of *E. coli* EPSPS fragment pairs consisting of the following: N67/C68, N85/C86, N104/C105, N154/C155, N182/C183, N184/C185, N218/C219, N224/C225, N227/C228, N259/C260, N298/C299, N371/C372, N376/C377, N383/C384, or is selected from a group of *P. putida* CGMCC 0739 EPSPS fragment pairs consisting of the following: N208/C209, N214/C215, N219/C220, N222/C223, N224/C225.
5. A nucleic acid molecule encoding the protein fragment according to any one of claims 1 to 4.
6. An expression vector harboring the nucleic acid molecule according to claim 5.
7. A cell comprising the nucleic acid molecule according to claim 5, or the expression vector according to claim 6.
8. A method for reconstitution of EPSPS, which comprises reconstituting EPSPS activity by using protein fragments according to any one of claims 1 to 4, or nucleic acid molecules according to claim 5, or expression vectors according to claim 6 in the absence of any joint structure.
9. A method for dividing EPSPS or EPSPS nucleic acid molecule, which comprises dividing the EPSPS into protein fragments according to any one of claims 1 to 4, or dividing the EPSPS nucleic acid molecule into the nucleic acid molecule according to claim 5 in the absence of any joint structure.

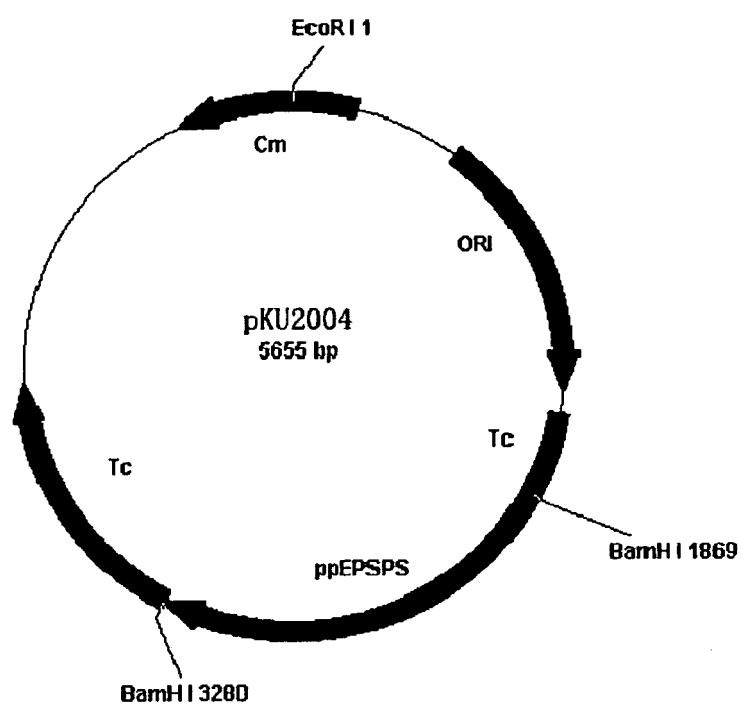


FIG. 1

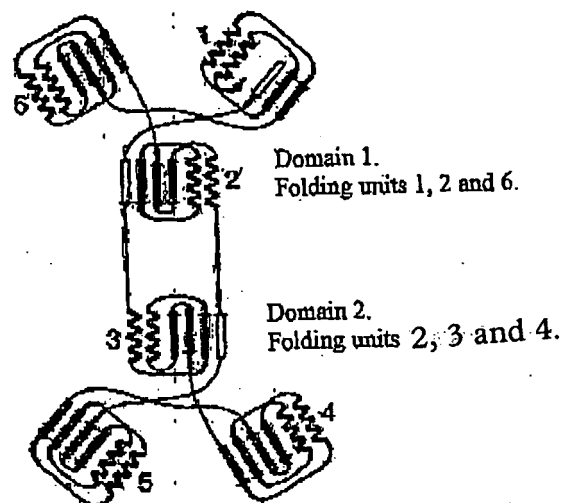


FIG. 2A



FIG. 2B

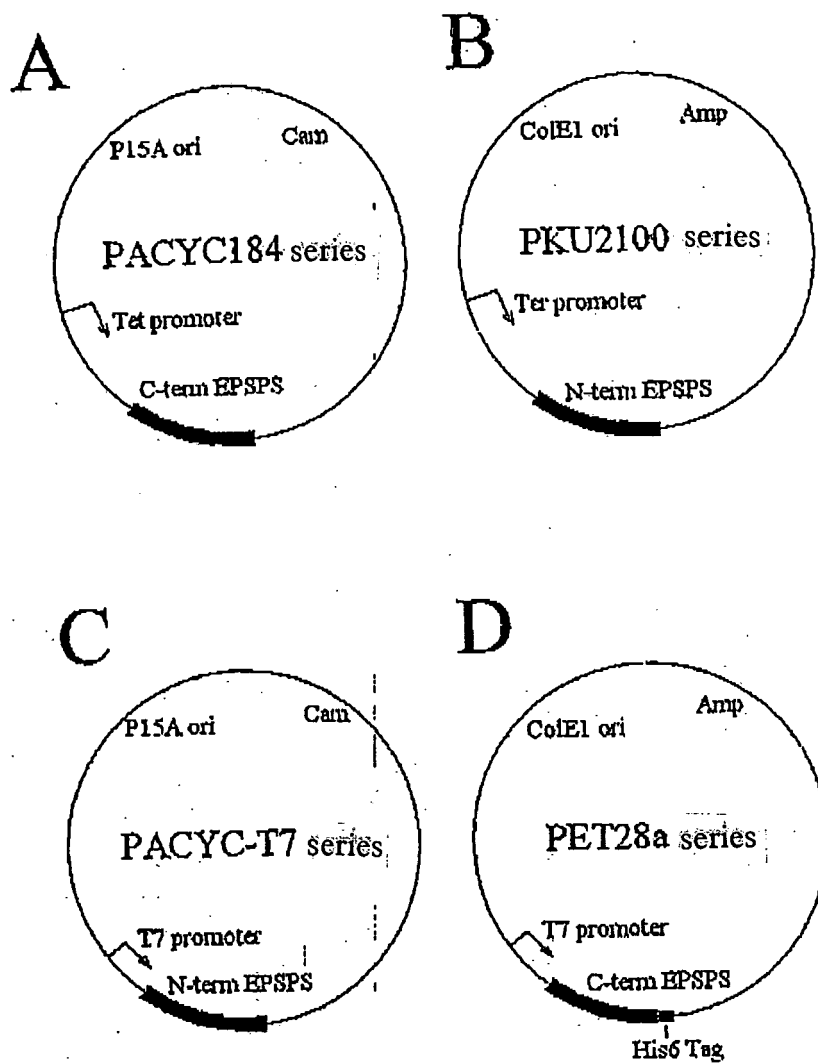


FIG. 3

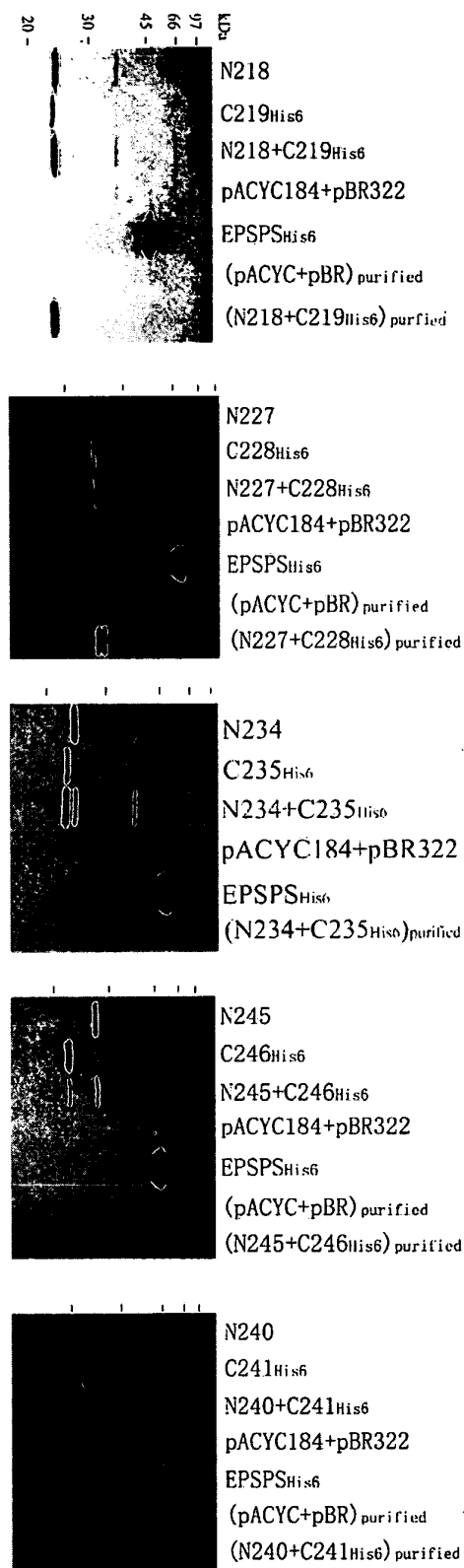


FIG. 4

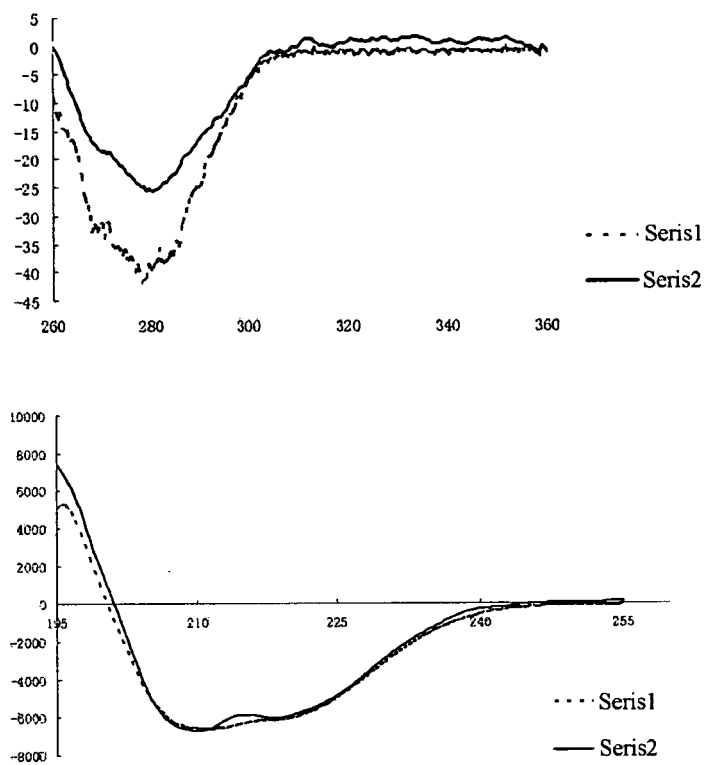


FIG. 5

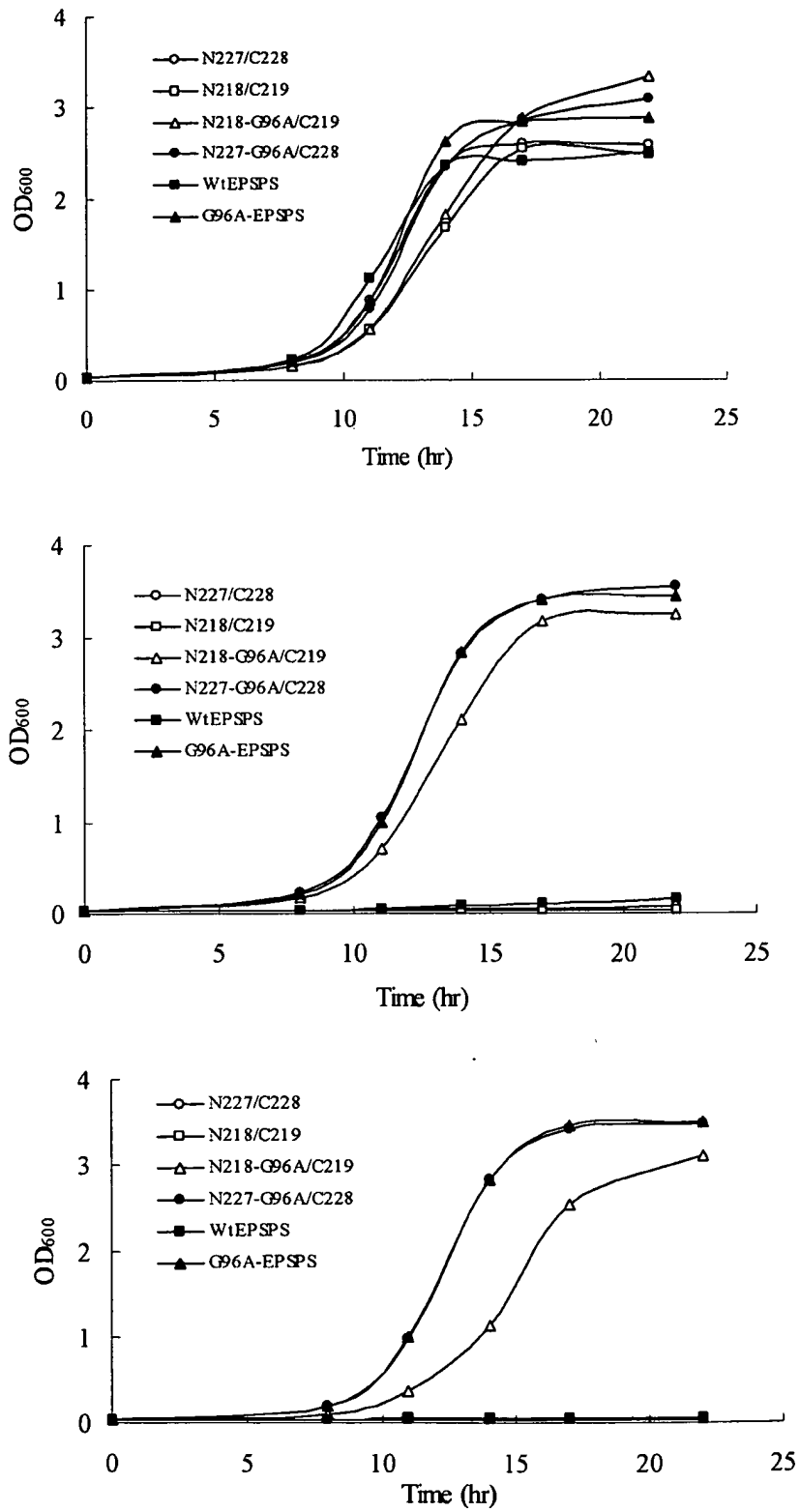


FIG. 6

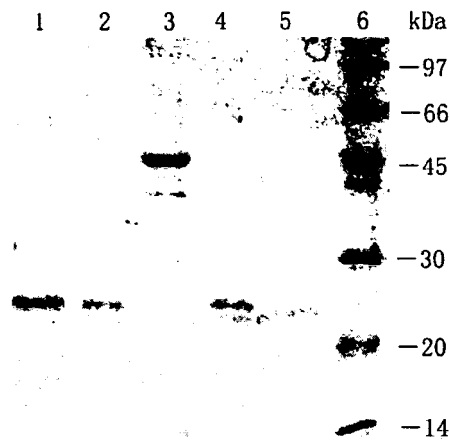


FIG. 7a

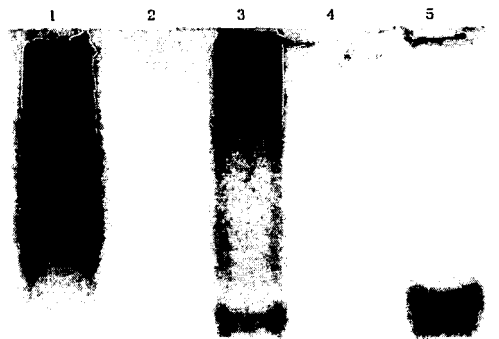


FIG. 7b

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2005/002219

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC⁸: C12N, C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

CNPAT, CNKI

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, PAJ, MEDLINE, BA, GENBANK (enolpyruvylshikimate, phosphate, EPSPS, AroA, glyphosate, herbicide, split-transgene containment)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Chen L et al. "Herbicide resistance from a divided EPSPS protein: the split Synechocystis DnaE intein as an in vivo affinity domain", Gene, vol. 263, no.1-2, Jan, 2001, pages 39-48, see the whole document	1-9
A	WO,A1,0071701 (NEW ENGLAND BIOLABS, INC.) 30.Nov 2000 (30.11.2000), see the whole document	1-9
A	Chin H G et al. "Protein trans-splicing in transgenic plant chloroplast: Reconstruction of herbicide resistance from split genes", PNAS, vol.100, no.8, Apr 15, 2003, pages 4510-4515, see the whole document	1-9

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
02.Mar. 2006 (02.03.2006)Date of mailing of the international search report
23 · MAR 2006 (23 · 03 · 2006)Name and mailing address of the ISA/CN
The State Intellectual Property Office, the P.R.China
6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China
100088
Facsimile No. 86-10-62019451

Authorized officer

LIN, Junkai

Telephone No. (86-10) 62085096



INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2005/002219

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Ye G N et al. "Plastid-expressed 5-enolpyruvylshikimate-3-phosphate synthase genes provide high level glyphosate tolerance in tobacco", Plant J, vol.25, no.3, 2001, pages 261-270, see the whole document	1-9

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2005/002219

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
WO,A1,0071701	30.11.2000	AU,B2,780002	24.02.2005
		AU,A,0050395	12.12.2000
		EP,A1,1183346	06.03.2002
		CN,A,1350582	22.05.2002
		JP,T,2003505012T	12.02.2003
		US,A1,2004096938	20.05.2004
		US,B1,6858775	22.02.2005

Form PCT/ISA /210 (patent family annex) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2005/002219

CLASSIFICATION OF SUBJECT MATTER

C12N 9/10 (2006.01) i
C12N 15/54 (2006.01) i
C12N 15/63 (2006.01) i
C12N 15/82 (2006.01) i
C12P 21/02 (2006.01) i

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- CN 02117991 [0008]
- WO 2004046359 A [0031]
- WO 2004046360 A [0031]

Non-patent literature cited in the description

- **WRUAD et al.** *Proc Natl Acad Sci U S A*, 2004, vol. 101 (40), 14533-8 [0002]
- **DANIELL et al.** *Nat Biotechnol*, 1998, vol. 16 (4), 345-8 [0003]
- **YE, G. N. et al.** *Plant J*, 2001, vol. 25 (3), 261-70 [0004]
- **CHIN, H. G. et al.** *Proc Natl Acad Sci U S A*, 2003, vol. 100 (8), 4510-5 [0004]
- **HAKANSSON, M. et al.** *Curr Protein Pept Sci*, 2002, vol. 3 (6), 629-42 [0005]
- **BRAUN, M et al.** *J Bacteriol*, 2003, vol. 185 (18), 5508-18 [0005]
- **SHIBA, K. et al.** *Proc Natl Acad Sci USA*, 1992, vol. 89 (5), 1880-4 [0005]
- **SHIBA, K. et al.** *J Biol Chem*, 1992, vol. 267 (32), 22703-6 [0005]
- **NELSON, K. E. et al.** Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ Microbiol*, 2002, vol. 4 (12), 799-808 [0005]
- **STALLINGS, W. C. et al.** *Proc Natl Acad Sci USA*, 1991, vol. 88 (11), 5046-50 [0016]
- **SCHONBRUNN, E. et al.** *Proc Natl Acad Sci U S A*, 2001, vol. 98 (4), 1376-80 [0017]
- **MCDOWELL et al.** *J Biomol NMR*, 2004, vol. 28 (1), 11-29 [0017]
- **ANDERSON, K. S. et al.** *JBiol Chem*, 1990, vol. 265 (10), 5567-72 [0018]
- **HUYNH, Q. K. et al.** *J Biol Chem*, 1988, vol. 263 (24), 11636-9 [0018]
- **HUYNH, Q. K. et al.** *J Biol Chem*, 1988, vol. 263 (2), 735-9 [0018]
- **PADGETTE, S. R. et al.** *J Biol Chem*, 1988, vol. 263 (4), 1798-802 [0018]
- **PADGETTE, S. R. et al.** *Arch Biochem Biophys*, 1988, vol. 266 (1), 254-62 [0018]
- **ESCHENBURG, S. et al.** *J Biol Chem*, 2003, vol. 278 (49), 49215-22 [0018]
- **MIZYED, S. et al.** *Biochemistry*, 2003, vol. 42 (23), 6986-95 [0018]
- **SHUTTLEWORTH, W. A. et al.** *Biochemistry*, 1994, vol. 33 (23), 7062-8 [0018]
- **SHUTTLEWORTH, W. A. et al.** *Arch Biochem Biophys*, 1996, vol. 334 (1), 37-42 [0018]
- **SHUTTLEWORTH, W. A. et al.** *Biochemistry*, 1999, vol. 38 (1), 296-302 [0018]
- **STAUFFER, M. E. et al.** *Biochemistry*, 2001, vol. 40 (13), 3951-7 [0018]
- **STAUFFER, M. E. et al.** *FEBS Lett*, 2001, vol. 499 (1-2), 182-6 [0018]
- **MCDOWELL, L. M. et al.** *Biochemistry*, 2004, vol. 43 (21), 6606-11 [0018]
- **SCHONBRUNN et al.** *Proteins*, 2000, vol. 40 (2), 290-8 [0019]
- *Biochemistry*, 2000, vol. 39 (9), 2164-73 [0019]
- *Proc Natl Acad Sci USA*, 2000, vol. 97 (12), 6345-9 [0019]
- **BICKEL et al.** *Phytochemistry*, 1978, vol. 17, 119-124 [0030]
- **MOUSDALE et al.** *Planta*, 1987, vol. 170 [0030]
- *Plant Physiol.*, 1987, vol. 83 [0030]
- *J Biol Chem*, 15 October 1988, vol. 263 (29), 15104-9 [0030]
- *Mol Gen Genet*, 01 December 1994, vol. 245 (5), 616-22 [0030]
- *Mol Gen Genet*, June 1993, vol. 239 (3), 416-24 [0030]
- **HANAHAN D.** *J Mol Biol*, 1983, vol. 166, 557-80 [0033] [0069]
- **CHANG, A. C. et al.** *J Bacteriol*, 1978, vol. 134, 1141-56 [0033] [0069]
- **SCHAEFFER, F. et al.** *EMBO J.*, 1982, vol. 1, 99-105 [0033] [0069]
- **LINK, A.J. et al.** *J. Bacteriology*, 1997, vol. 179, 6228-6237 [0033]
- *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 1989 [0038]